# INDIVIDUAL DIFFERENCES IN IMPLICIT LEARNING: INITIAL EXPLORATION OF CLINICAL UTILITY

A Thesis

Submitted to the Faculty

of

Drexel University

by

Daniel Grady Smith

in partial fulfillment of the

requirements for the degree

of

Master of Science in Psychology

December 2013

## Acknowledgments

I am greatly appreciative for the support of my thesis committee: Felicia Hurewitz, Ph.D., Mary Spiers, Ph.D., and John Herrington, Ph.D. Their continued support through the long process leading to the present proposal has provided direction through chaos, and strength through struggle. I am also grateful for the hard work of my colleagues for their various roles in the development of this project: Nicole Roshannon, Jennifer Mongiat, Caitlin Seldin, Megan Hartz, and Evelyn Hinkle. I must acknowledge the crucial role Cory Czuczman has played in putting the 1s and 0s together to make our cognitive learning tasks fun. I am additionally appreciative of the end-stage programming of Dilan Alpay and Ethan Miller and of the art and animation provided by Terry DiFeliciantonio and Valentina Feldman.

I am eternally grateful to my family, Esther Ellison and Selah Bjornson-Smith for their constant laughs, stubborn determination, and love; I could not have gotten here without you both.

## Table of Contents

## Abstract

Individual Differences in Implicit-Statistical Learning: An Initial Exploration of Clinical Utility
Daniel Grady Smith
Felicia Hurewitz, Ph.D.

At present, no reliable implicit learning (IL) measure exists for use in the clinical setting, though recent evidence suggests that IL can predict individual differences in syntactic comprehension and word segmentation among and between neuropsychiatrically typical and atypical individuals. In the present series of studies, we designed and tested an adaptation of a paradigmatic IL task (*Serial Response Time Task*) with the goal of creating a motivating and practical clinical tool. The pilot and primary experiments (*N* = 7, *N* = 41, respectively) revealed that reliable individual differences in IL are observed during our versions of the task, but only when accounting for confounds related to both the probability and magnitude of movement in a mixed-effects model analysis. Correlational analyses from a subset of individuals completing the task (*N* = 24) indicate that implicit learning was not associated with syntactic comprehension, working memory, or IQ. We conclude that this new task is promising given it is a substantially shorter task than those typically used in the experimental literature, yet is nevertheless able to detect reliable individual differences despite confounds. Several modifications to the task are recommended that aim to reduce the influence of confounds, and more cleanly represent IL ability.

## Chapter 1: Introduction

Implicit learning represents the process of analyzing the statistical properties of stimulus distributions over time allowing for the identification of associations (Misyak, Goldstein, and Christiansen, 2007) with a minimum of required attention and awareness (Shanks, 2005; Perruchet and Pacton, 2006). Recent evidence suggests that individual differences in implicit knowledge acquisition may be meaningfully associated with higher-order cognitive and adaptive functions. Specifically, implicit learning has been associated with a variety of crucial tasks such as language acquisition (c.f., Reber, 1965; Kuhl et al., 1992; Saffran, Aslin, and Newport, 1996; Saffran, 2001; Misyak, Goldstein, and Christiansen, 2007), learning complex movement patterns (Nissen and Bullemer, 1987; Jimenez and Vazquez, 2005), and acquired knowledge of social conventions (c.f., Lieberman, 2000). It is therefore unsurprising that clinical populations with impairments in these higher-order processes have been shown to have deficits in implicit learning (e.g., Parkinson's Disease, Knowlton, 2002; Dyslexia, Folia et al., 2008; Autism Spectrum Disorders, Mostofsky et al., 2000; Specific Language Impairment, Evans, Saffran, and Robe-Torres., 2009). Given these strong empirical foundations, it can be argued that normative implicit learning tasks based upon those tasks used most frequently in the literature might prove invaluable in clinical assessment. However, no attempts exist to develop a neuropsychological assessment of implicit learning based upon these widely used tasks.

In a first attempt to extend the already robust research literature into clinical practice, we adapt a paradigmatic task from the implicit learning literature for use with a variety of clinical populations. Specifically, these adaptations attempt to make these computerized tasks usable across a variety of developmental levels as well as appropriate for populations having a wide range of intellectual and adaptive ability.

### *1.2. Implicit Learning: Definition and Task Background*

The implicit learning literature represents an endeavor to characterize a species-general ability to recognize associated events and patterns in the environment. Many theories posit primarily unconscious and unintentional aspects as opposed to deliberate, intentional processes. To use formal definitions within this literature, the unconscious and unintentional learning system – the *implicit learning system* – is capable of recognizing patterns in the environment without the actor's awareness that such recognition, or analyses, of patterns are taking place (Shanks, 2005). Nevertheless, the learner is able to utilize this information to adapt their behavior by utilizing these patterns. In contrast, the *explicit learning system* is characterized by the conscious, deliberate and intentional effort to gain knowledge of patterns within the environment (c.f., Shanks, 2005).

The distinction between the unconscious and the conscious remains contentious as several researchers clearly delineate a separate cognitive and neurological component for each system, while others argue for a single, unified theory of learning (see Mitchell et al., 2009 and subsequent commentary for a review of this debate). For the present thesis, we take the dual-systems approach, maintaining that a separate implicit learning construct with a distinct set of processing mechanisms exists and is worthy of investigation. We focus on implicit learning, because clinically relevant aspects of this unconscious learning process, such as meaningful individual differences, are only recently beginning to be understood. Several paradigms (and their modifications) have accounted for a great deal of present understanding of the ability to acquire knowledge of environmental patterns implicitly and so will be reviewed.

### *1.3. Artificial Grammar Learning Task*

Reber first formally investigated implicit learning in response to claims that humans were capable of acquiring knowledge of the grammatical rules in their language in a non-conscious and unintentional manner (c.f., Reber, 1967). Prior to this, the nature of human implicit learning had not been investigated, except for in the perceptual learning literature by Gibson & Gibson (1955). As participants seemed able to "respond to the statistical nature of [a] stimulus array" (Reber, 1967, pp. 856) without awareness, an investigation into the learning mechanisms underlying such abilities was imperative to further develop theories of unconscious acquisition of knowledge.

To explore this issue further, Reber (1967) developed the Artificial Grammar Learning (AGL) task. In AGL paradigms, a finite-state grammar is first determined. Briefly, a finite-state grammar contains a set number of internal states, with movement from one state to the next resulting in specific final states. Once the finite-state grammar is created, specific rules are formed dictating the way by which a sequence, from the initial- to end-states may be constructed. The task is presented in detail below as it provides the foundation of many other AGL paradigms and investigations into implicit learning processes. Thus an understanding of Reber's groundbreaking paradigm will prepare the reader to understand many of the arguments presented by his contemporaries.

Reber's (1967) AGL paradigm was constructed with five letters (P, S, T, V, X), as shown in the schematic representation in figure 1. The specific sequences in the experiment were restricted to lengths of 3 – 8 items. For example, Reber's artificial grammar could produce TPPTXXVS, by following the red path through figure 1. Briefly, the red path indicates "T", "P" with a recursion resulting in "PP", then "T" as it is the only available option, "X" diagonally then an "X" without a recursion, "V", and finally "S", resulting in the output ("TPPTXXVS"). During the experiment, the participants were first exposed to a learning phase consisting of 20 trials. During the learning phase they were required to remember the sequences of letters briefly but given no information about the underlying grammatical structure. During a later testing phase, participants were told that the items actually had a complex grammatical relationship and that they were to make grammaticality judgments (i.e., is the sequence grammatical or not). Reber's analysis of these data indicated, not only that participants could accurately make grammatical judgments, but that these judgments revealed robust implicit learning mechanisms as participants admitted being unaware of the underlying structure. Specifically, Reber's experiment indicated that the lawfulness of stimulus sequences could be abstracted from the environment without the involvement of explicit, verbalizable, strategies. In the almost five decades since Reber's initial investigation, there have been a great many

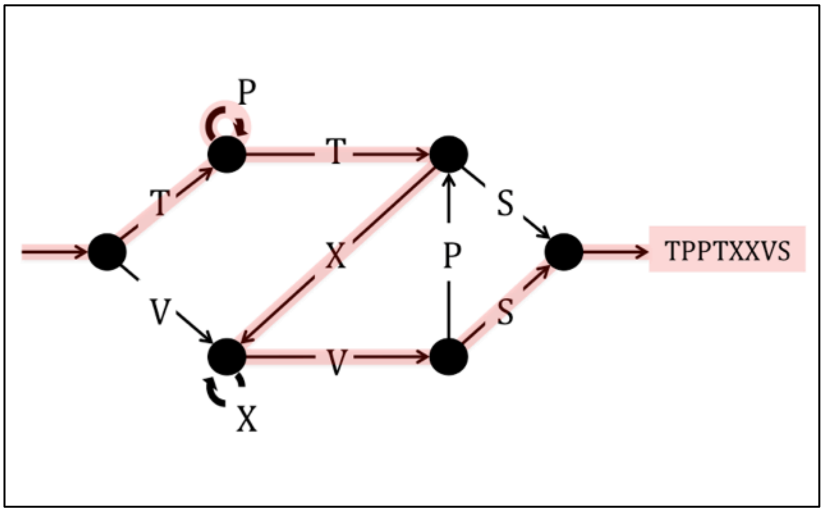


***Figure 1. Artificial Grammar Learning task as used by Reber (1967).*** The above schematic shows a grammatical structure indicating possible sequences of letters. Sequences start on the left and proceed through each decision point (circles) along several possible paths until they reach the final decision point on the right (at which point the sequence is outputted). The red highlighting denotes a possible path leading to the 8-item sequence *TPPTXXVS*.

permutations of the initial AGL paradigm, but much of the theoretical utility underlying the procedures remain the same (see Pothos, 2007 for a review).

Although the previous evidence provided by the AGL paradigm indicated acquisition of complex information was possible without explicit instructions to learn, the role of awareness and attention in implicit learning had not yet been established. While it was true that Reber did not give the participants explicit instructions to memorize the grammatical rules, this does not preclude the participants from consciously extracting some of the underlying patterns.

To address the role of attention in memory acquisition directly, Nissen and Bullemer (1987) had healthy controls participate in both single- and dual-task conditions as well as compared their performance to an amnesic group, known to have memory encoding deficits.

*1.4. Serial Response Time Task*

Nissen and Bullemer's participants completed an *online* (in vivo) measure of implicit learning, the serial reaction time task (SRTT). The SRTT is an *online measure* in the sense that acquisition of skill or knowledge was measured as learning occurred. This stood in contrast to other implicit learning tasks of the time (including the AGL described above) in that many measured the end state of learning through the facilitation effects on motor performance or recall and recognition of specific material (e.g., priming paradigms such as word fragment completion; c.f., Warrington and Weikrantz, 1974).

During a typical SRTT paradigm, participants are required to respond to stimuli on a computer screen with a corresponding button press on a keyboard or response box. Unknown to the participants, the stimuli appear on the screen in a repetitive sequential order during many of the presentations. Nissen and Bullemer's paradigm presented participants with an asterisk in one of four locations on a computer screen one at a time and the participant was to press a corresponding key on the keyboard that was specially marked (see figure 2).

On a given trial, participants see a stimulus in one of several predetermined positions. In the case of many experiments, these are represented by four horizontal locations along a computer screen's *x-axis*. In our task, we divided an iPad screen (in landscape orientation) into four equally spaced columns, allowing stimuli to appear in one of the columns on each trial. As soon as the participant sees a stimulus the participants had to touch the stimulus on the screen directly.

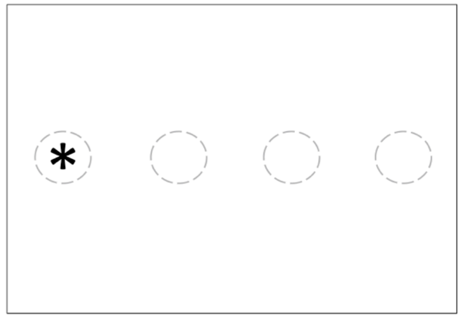

***Figure 2. Typical Serial Response Time Task presentation.*** Participants see stimuli (such as asterisks) in one of four spatial locations (dashed circles denoting possible positions). Participants are instructed to press a corresponding key on a keyboard or button box as quickly as they can. Certain blocks of trials typically have a specified order (e.g., 4, 2, 3, 1, 3, 2, 4, 3, 2, 1), while other blocks are randomized.

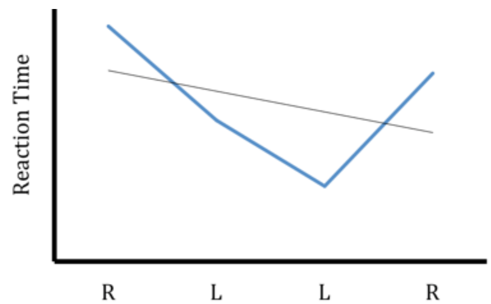


***Figure 3. Representative results of learning on the SRTT.*** In four block SRTTs (e.g., present experiment; Mostofsky et al., 2000), the first and fourth block contain randomized sequences (*R* blocks), while the second and third contain many repetitions of the same sequence (*L* blocks). Typically , participants become faster to all blocks across time as they acclimate to the task (represented by the black linear trend line). However, they perform significantly faster over the *L* blocks compared to *R* blocks. Additionally, when the second *R* block is presented, performance slows significantly compared to the second L block. This is thought to indicate impaired performance after deviation from an implicitly learned sequence (Shanks, 2005).

In Nissen and Bullemer's (1987) task, each participant completed two conditions: a learning condition consisting of 10 repetitions of a 10-item sequence (e.g., position sequence: 4, 2, 3, 1, 3, 2, 4, 3, 2, 1); or a random condition in which the location on each trial was random across 100 trials. The results revealed (and have many iterations of the paradigm since; see Shanks, 2005) that individuals learn the sequence within the learning condition even though they were not told sequential information within the trials provided a cue. That is, performance on the SRTT paradigm decreases with repeated exposure to sequential trials to a far greater degree than trials following a random presentation sequence (for an example see figure 3).

Nissen and Bullemer's comparison between individuals with Korsakoff's syndrome and neurologically healthy controls indicated that awareness was not required to learn. That is, individuals with Korsakoff's syndrome were able to learn the visuomotor sequences despite their known deficits in "consciously recalling verbal and nonverbal information" (Nissen and Bullemer, 1987, pp. 30). In further exploration, Jimenez and Vazquez (2005) provide evidence that SRTT experiments utilizing deterministic sequences, such as those used by Nissen and Bullemer (1987) will be likely to engage both explicit and implicit processes. When explicit processes are activated, so too are central attentional resources. However, by modifying the task to include a probabilistic sequence (less noticeable to many participants) competition for attentional resources (e.g., dual-task conditions) has little effect on learning.

In further exploration, Jimenez and Vazquez (2005) provide evidence that SRTT experiments utilizing deterministic sequences, such as those used by Nissen and Bullemer (1987) will be likely to engage both explicit and implicit processes. When explicit processes are activated, so too are central attentional resources. However, by modifying the task to include a probabilistic sequence (less noticeable to many participants) competition for attentional resources (e.g., dual-task conditions) has little effect on learning.

As has been seen in the review thus far, learners are able to acquire abstract knowledge of the grammatical (i.e., rule-based) and sequential relationships between stimuli through implicit learning mechanisms. Until relatively recently, such knowledge of stimulus associations was thought to be reliant upon relatively simplistic analyses, as complex statistical computations of stimulus distribution information were, "generally considered to be too complex for [learners] to use" (Saffran, Newport, and Aslin, 1996, pp. 608). However, research by Saffran and colleagues (e.g., Saffran et al., 1996; Saffran et al., 1997; Aslin et al., 1998; Saffran, 2002), indicated that in fact, human learners as young as 8 months of age are capable of computing a range of complicated statistics and furthermore, that they might use these complicated computations to learn certain aspects of language.

### *1.5. Summary of Implicit Learning Paradigms*

Taken together, the heavily investigated AGL and SRTT tasks provide crucial insight into the implicit and incidental computational mechanisms available to learners. As the mechanisms underlying learning on these tasks are thought to be similar, recent research has explored the relationship between a participant's performance on one task with performance on the other, even though the stimuli require the processing of disparate forms of information (e.g., visuomotor sequential vs. abstract grammar). From this perspective, individual differences in one task may be seen as predicting a domain general implicit learning ability. Thus correlations among tasks requiring the use of this ability are expected. However positing implicit learning as an ability is surprising as early theory argued for limited variability in implicit learning (Reber, 1991). Nevertheless several very recent paradigms indicate that implicit learning ability may play an important contributory rule in many aspects of cognitive function above and beyond those described in general intelligence (*g*) defined previously (e.g., Carrol, 1993). Similarly, implicit learning ability may play an important role in the deficits observed in several neurocognitive and developmental disorders. It is to these issues we turn next.

## Chapter 2. Individual Differences in Implicit Learning

### *2.1 Historical Foundations of the Individual Differences Argument*

According to Reber and colleagues (1989, 1991) implicit learning processes should be “viewed within the context of adaptionist principals of evolutionary biology. . . in terms of ontogeny, phylogeny, and function” (Reber, Walkenfeld, and Hernstadt, 1991, pp. 888). As such Reber and colleagues argued for the “primacy of the implicit,” indicating further that implicit learning represents “the functional instantiations of a phylogenetically primitive system that developed before the emergence of conscious functioning” (Reber, Walkenfeld, and Hernstadt, 1991, pp. 888). Given that implicit learning processes were hypothesized to be phylogenetically older, they were thought to have certain characteristics that would set them apart from conscious explicit processes (e.g., problem solving). Specifically, Reber, Walkenfeld, and Hernstadt argued (pp. 888):

1) [sic] Implicit systems should be robust in the face of various psychiatric or neurological insults which compromise functioning of the [explicit processes].
2) Implicit systems ought to display tighter distributions [and thus fewer individual differences than explicit systems].

3) Implicit functions should operate largely independently of standard measures of cognitive capability such as intelligence, assuming intelligence is being measured by [an] IQ test.

In regard to the first argument, Reber and colleagues provide a comprehensive (at the time) review of the literature indicating that deficits in implicit learning processes had not appeared. This was despite investigations of implicit learning ability in syndromes involving the memory structures of the medial temporal lobe (Korsakoff's syndrome, Alzheimer's disease) as well as neuropsychiatric disorders such as schizophrenia and depression (Reber, Walkenfeld, and Hernstadt, 1991). In regard to the second argument, Reber and colleagues' review indicated no reported individual differences in implicit memory when compared against SAT, age, or school performance. The third argument however, had not been previously tested, and was thus investigated directly for the first time by Reber, Walkenfeld, and Hernstadt (1991). Specifically, Reber and colleagues compared performance on the AGL to performance of a primarily explicit task as well as to a measure of intelligence (WAIS-R). The results indicated far fewer individual differences in the implicit task as variance was much smaller than that observed in the explicit task. Furthermore, the results indicated a strong correlation between explicit task performance and IQ but an insignificant correlation between implicit performance and IQ.

However, Mackintosh (1998), argues that standardized intelligence tests primarily take into account explicit learning and processing ability. Mackintosh further argues that as the implicit learning system plays a crucial role in non-consciously detecting environmental contingencies, individual differences in implicit learning should indeed be taken into account despite implications of Reber's position. Specifically, Mackintosh argues that individual differences in implicit learning do have meaningful associations with real-world performance, even though the variance might be tighter in implicit learning tasks, are unimpaired in several disorders, and do not have associations with intelligence as measured by IQ (c.f., Mackintosh, 1998). Until recently however, Mackintosh's position lacked empirical evidence as the association between implicit learning processes and cognitive ability had not been tested beyond associations with standardized IQ tests (Reber, Walkenfeld, and Hernstadt, 1991; Gebauer and Mackintosh, 2007).

### *2.2. implicit learning and General Cognitive Ability: States Versus Traits*

In an attempt to determine the psychometric properties of general intelligence, implicit learning measures, and dynamic decision-making, Danner et al. (2011), compared each construct within a latent state-trait theory framework. Such analyses are able to provide a delineation of states (ephemeral, time specific performances) and traits (persistent performances thought to represent underlying constructs). Such a framework requires measurement to be taken on two occasions to determine the stability of the measure across time. Implicit learning ability and dynamic decision-making were additionally compared to measures of professional success to determine their contribution above and beyond IQ.

Danner and colleagues (2011) included 173 individuals with 151 completing the second measurement. Each individual completed two measures of intelligence (Raven's Advanced Progressive Matrices: Raven, Court, and Raven, 1994; Berlin

Intelligence Structure Test: Jager, Sub, and Beauducel, 1997), one implicit learning task (AGL described above), two dynamic decision making tasks (Heidelberg finite state automaton: Wirth and Flunke, 2005; Tailorshop: Funke, 1983), and a measures of objective professional success (income, self-rated social status, and highest educational attainment). Results of the latent state-trait analysis indicated that cognitive measures were indeed unassociated with situational factors such as fatigue (a first for the AGL paradigm) and were thus more likely to represent traits. Additionally, implicit learning and dynamic decision making tasks were significantly correlated with intelligence, objective measures of professional success, and with each other. Together these results indicate that dynamic decision making, and importantly for the present discussion, implicit learning were shown to both represent measurable traits as well as predict professional success. However, in comparing trait specificities, the medium size between implicit learning and professional success appeared to be accounted for by IQ, suggesting no predictive value for implicit learning above and beyond intelligence. Danner and colleagues indicate however, that this low trait-specificity may have been associated with unsystematic measurement error. Thus, while Danner and colleagues' results are promising, they admit they are unable to measure implicit learning in a reliable way using the AGL paradigm. Therefore the meaningfulness of individual differences in implicit learning must remain undecided based on this study.

*2.3. Implicit learning and General Cognitive Ability: Implicit learning is Separate but Important*

In order to provide a comprehensive examination of meaningful individual differences in implicit learning, Kaufman and colleagues (2010) used structural equation modeling to determine the relationships among a large battery of cognitive tasks including a probabilistic SRTT. The probabilistic SRTT proceeds in the same manner as the SRTT described previously, only that control (random) trials are interspersed within the sequence in such a way as to maintain a specific probability that a particular sequence will occur. This task was thought to more clearly match real world implicit learning, as learning in the natural environment is more likely to encounter noisy and probabilistic distributions of stimuli as compared to deterministic sequences (Jimenez and Vazquez, 2005).

Although somewhat similar in scope to the work by Danner et al. (2011), the use of the SRTT paradigm in the Kaufman study provides several advantages. Specifically the SRTT paradigm has been argued to be a distinctly better measure of implicit learning than the AGL paradigm used by Danner and colleagues (Destrebecqz and Cleeremans, 2001). As indicated by Kaufman and colleagues (2010), SRTT paradigms are more likely to result in incidental acquisition as individuals are simply responding to stimuli, whereas in AGL paradigms they are asked to explicitly memorize information. This could potentially result in explicit processing of the stimulus array. Additionally, SRTT measures provide an online measure of learning based upon reaction time, whereas AGL paradigms require an explicit retrieval process during a testing phase, which is wholly separate from the learning phase. The aforementioned measurement error seen in the AGL task by Danner and colleagues may be less likely to occur in the probabilistic SRTT paradigm used by Kaufman et al. (2010). Thus, it can be argued that the structural equation model used by Kaufman and colleagues, provides stronger evidence that meaningful individual differences exist in implicit learning.

In their structural equation model, Kaufman et al., included measures of implicit learning (probabilistic SRTT), psychometric intelligence, working memory, explicit associative learning, processing speed, academic achievement, and personality. The results from the cognitive tasks alone indicate that the latent variable of implicit learning shares a small but significant relationship with indices of processing speed and verbal reasoning, but non-significant relationship with each other cognitive latent variable (psychometric intelligence, working memory, explicit learning). Additionally, implicit learning was significantly correlated with math and foreign language achievement.

Taken together, Kaufman and colleagues results replicate the previously reported findings by Reber (1991) in that individual differences in implicit learning ability were not associated with psychometric intelligence. The findings additionally indicate a link between processing speed and implicit learning, which the authors take to indicate that perhaps both processing speed and implicit learning (as per Reber et al., 1991) are "phylogenetically older" cognitive constructs (Kaufman et al., 2010, pp. 335). In contrast to the aforementioned support of the theoretical positions by Reber and colleagues (1991), the argument positing lower variability in implicit learning should be qualified. Specifically, Kaufman and colleagues find that even if variability is low, individual differences in implicit learning are shown to have meaningful associations with multiple aspects of human behavior.

Of particular note is the association between implicit learning and indices of second language acquisition as it appears to be that individual differences in implicit learning learning mechanisms are able to account for a meaningful proportion of the variance in the observed indices of language acquisition.

### *2.4. Individual Differences in Implicit learning and Language*

The work by Kaufman et al., (2010) argues for an association between domain general implicit learning and language acquisition as language specific implicit learning mechanisms were not tested. Thus any underlying association must be attributable to some shared mechanism between visuomotor sequence learning on the one hand (probabilistic SRTT) and language acquisition and reasoning on the other.

Similarly, Misyak and colleagues utilize more clearly focused comparisons across a number of studies and continue to find a primary role for domain general implicit learning in language comprehension (Misyak and Christiansen, 2007, 2012; Misyak, Christiansen, and Tomblin, 2010a, 2010b). In their most recent study, Misyak and Christiansen (2012) compare three statistical learning tasks to a measure of syntactic comprehension. As the authors seek to determine the unique contribution of implicit learning to syntactic comprehension, they additionally include in their model measures of lexical and cognitive abilities known to be associated with language comprehension (i.e., vocabulary, reading experience, working and short-term memory, fluid intelligence). The results of their study indicate that language specific implicit learning does indeed predict language comprehension above and beyond the influence of other cognitive and lexical factors. This argues for a unique and important role of implicit learning in language processing abilities.

### *2.5. Summary*

Reber and colleagues' argument that implicit learning should be separate from measures of psychometric intelligence is supported by a variety of well-supported experimental findings (e.g., Kaufman et al., 2010; Misyak and Christiansen, 2007). However, Reber and colleagues indication of low expected variability in implicit learning seems to underplay the importance of individual differences in implicit learning observed to correlate with several important cognitive skills (e.g., syntax acquisition, Kidd, 2012; language comprehension, Misyak and Christiansen, 2012). Given this very recent "turning of the tide" in regards to the variability of human implicit learning, further investigation of the clinical utility of normative measures of implicit learning seem even more fruitful. Particularly those measures previously shown to exhibit meaningful individual differences associated with important adaptive functions (e.g., association between SRTT, AGL, and language as reported in this section). However, the review thus far has not addressed Reber and colleagues first and likely most relevant argument to the present thesis: implicit learning should be robust in the face of psychiatric or neurological insults. A breadth of information exists on this topic in the experimental literature. Rather than summarize the entirety of this literature, select, very active, programs of research will be discussed which focus on assessing the implicit learning abilities of individuals with an Autism Spectrum Disorder.

*2.6. Implicit Learning in Autism Spectrum Disorders*

There is growing interest in assessing the implicit learning abilities of individuals with autism as deficits in implicit learning provide a parsimonious explanation of the underlying etiology of observed deficits in in the disorder (Klinger, Klinger, and Pohlig, 2007). Eigsti presents a thorough description of implicit learning abilities in autism (Eigsti, 2011). Therein she briefly reviews qualitative evidence indicating that individuals with autism can be observed to have highly specialized skills in one area but typically fail to generalize (i.e., to gain knowledge of associations between stimuli or learning contexts that are not explicitly identified) (Klinger, Klinger, and Pohlig, 2007; Travers et al., 2010). These observations have baring on assumptions made about implicit processes as the ability to form prototypes, to categorize, and thus to create an abstract relationship from the specific to the generalized, is thought to rely in part on implicit learning (Eigsti, 2011). Importantly these abilities have been shown to be impaired in individuals with autism (Klinger and Dawson, 2001). It is also noteworthy that some of the anatomical regions involved in implicit learning (e.g., cerebellum, basal-ganglia, prefrontal cortex) show anatomical differences in autism (e.g., Corchesne, 2003). However, Mueller (2004), Happe (2006), and Geschwind (2007) indicate anatomical heterogeneity is characteristic of autism and thus identifying impaired regions via fMRI may hold for one group but not another.

Complex skill acquisition suggests another possible role for the involvement of implicit learning deficits in autism. Specifically, knowledge of sequential information, reviewed in depth previously, is involved in skills requiring complex movement sequences such as those used for social skills (e.g., waving, blowing a kiss; Eigsti, 2011), and non-verbal and verbal communication (Walenski, Tager-flusberg, & Ullman, 2006). There is also some indication that individuals with ASD have a tendency to rely upon explicit strategy formation as compared to implicit acquisition. Evidence here comes from Klinger and Dawson 2001 indicating that individuals with an ASD rely on explicit strategies when learning new info and categories. It has also been indicated that their word retrieval pattern is consistent with enhanced declarative memory involvement (Walenski et al., 2008).

Although there are a number of well-informed arguments for positing an implicit learning deficit in autism spectrum disorders, experiments directly investigating implicit learning in this population show mixed findings (for a review, see Eigsti, 2011). However, as individuals with ASDs express marked variability in cognitive functioning, Geschwind's theories concerning "the autisms" (2009) or dissimilar traceable phenotypic presentations, may best characterize this disorder. Thus, it may be the case that implicit learning deficits underlie a number of impaired abilities in autism, but only for some of the individuals affected. Furthermore, although the studies that have thus far investigated implicit learning ability in autism have used somewhat similar but divergent techniques, the comparisons are not based upon normative measures of implicit learning, as no such measures exist presently. Thus, the advantages of creating a normative implicit learning task for measuring meaningful individuals differences in cognitive functions (e.g., language comprehension) might be particularly useful in determining if implicit learning deficits exist in autism. In the present study, we investigated the initial utility of newly developed potential normative implicit learning measures. We next describe the development of these experimental tasks as well as describe another research team's previous attempt to create a clinical measure of implicit memory.

## Chapter 3: Implicit Learning Tasks Designed for the Clinical Setting

### *3.1. Lessons Learned: Previous Investigation of a Clinical Implicit Learning Measure*

Despite the aforementioned evidence indicating that meaningful individual differences exist and furthermore that specific deficits exist in this ability within certain populations, there has thus far not been any attempt (to our knowledge) to adapt paradigmatic implicit learning tasks (e.g., SRTT) for use as clinical instruments. There has however been one attempt to create a new normative test of implicit memory for use in clinical populations (Sopena et al., 2005; Kessels, Remmerswaal, and Wilson, 2011). However, the procedure utilized in the paradigm diverges greatly from a number of conventions in the implicit learning literature, and in turn, our discussion of the literature thus far. However, our motivation for developing normative implicit learning measures as well as the considerations we make in their development somewhat mirror those of Wilson and colleagues and we thus discuss their paradigm within this context.

In brief, the Implicit Memory Test (IMT) developed by Sopena and colleagues (2005) is a pen and paper task including both verbal (word stem completion) and visual (fragmented pictures) measures of implicit memory.  The use of pen and paper tasks in the IMT is a rather large break with tradition as many implicit learning measures include embedded computerized measures of reaction time. In-depth analyses of reaction time allow for fine-grained analyses of online (i.e., in vivo) learning or at the very least response time differences across learning conditions. However, as aptly noted by Wilson and colleagues, computerized measures are not necessarily easy to use in neurologically impaired populations "such as individuals with severe dementia" (Kessels, Remmerswaal, and Wilson, 2011, pp. 179) as they at times include a complicated setup as well as a large number of trials to evince learning. We agree with these contentions but would extend and update the logic underlying them.

Beyond the argument made by Kessels and colleagues (2011), it can be similarly argued that some incarnations of implicit learning tasks are not developmentally appropriate for pediatric patients, as they require prolonged attention to stimuli arguably uninteresting to many young participants. This might particularly be the case for individuals with an autism spectrum disorder and concomitant intellectual disability given the challenges in motivating this sub-population to perform at their best during neurocognitive assessments (Ozonoff, 2004).

Additionally, Kessels and colleagues (2011) argue that the of portability and setup are crucial factors to consider when designing neuropsychological assessment instruments, as assessments could take place at the bedside or generally away from the examiner's primary office space. While not a necessity, some implicit learning tasks have additionally used response boxes, or other apparatus, increasing the complexity of the setup and similarly decreasing their portability. For experimental purposes, this is quite acceptable, but for clinical measures this situation is less than ideal. Thus it seems an intuitive step to sacrifice the robustness yet complexity of computerized assessment for the simplicity and ease of travel provided by pen and paper measures as in the IMT. Laptops could have certainly been used, but again, specialized response boxes might have caused some difficulty.

In recent years however, tablets (such as the Apple iPad) have become far more prevalent and inexpensive. Benefitting experimenters, these forms of computing platforms provide a more direct and intuitive interaction between the user and the stimuli on the screen. Where a participant would have had to use a secondary input device such as a mouse or a keyboard to interact with objects, they can now directly interact with the virtual object by touching it on the screen. Direct interaction is a desirable quality for experimenters and clinicians as the level of explanation required for people unfamiliar with how to use a computer is minimal. This could be a particular advantage for research involving young children, older adults, and cognitively impaired populations. With these considerations in mind we created a touch screen task, adapted from the paradigmatic SRTT (Nissen and Bullemer, 1987) described previously.

## Chapter 4. Pilot Study: Initial Investigation of a New SRT Task

The SRTT paradigm was chosen as a candidate measure of clinically implicit learning as it provides an online, trial-by-trial measure of implicit learning allowing for in-depth analyses of individual learning as well as indicators of overall learning (i.e., learning condition compared to baseline) (Jimenez and Vazquez, 2005). The SRTT paradigm has additionally been shown to correlate with broader cognitive functioning (e.g., syntactic processing, Kidd, 2011). These qualities make the SRTT ideal for an exploration of meaningful individual differences in implicit learning.

In consideration of the complexity of SRTT paradigms and to make the tasks more appropriate for use in special populations, we decrease the required trials to four blocks of 72 trials. This represents a considerable decrease from typical SRTT paradigms as the number of required trials can at times exceed 500 (Jimenez and Vazquez, 2005). However, our total of 288 trials is still a rather large number. Nevertheless we hesitate to decrease the number of trials further given the complex nature of our adapted sequence of stimuli. To learn this sequence it is likely that a certain number of trials are required to evince learning. Therefore, an alternative to trial reduction was needed. To this end we attempted to make the task engaging and developmentally appropriate, by mimicking game-like qualities that add an arguably intuitive understanding of task goals. Specifically, rather than typical response box or keyboard press implementations of our SRTT task, virtual bubbles appeared on the iPad screen that were to be popped (see figure 4). To increase engagement in the task, the bubbles would pop and be accompanied by a popping sound, thought to represent an auditory reward for interacting with the stimulus.

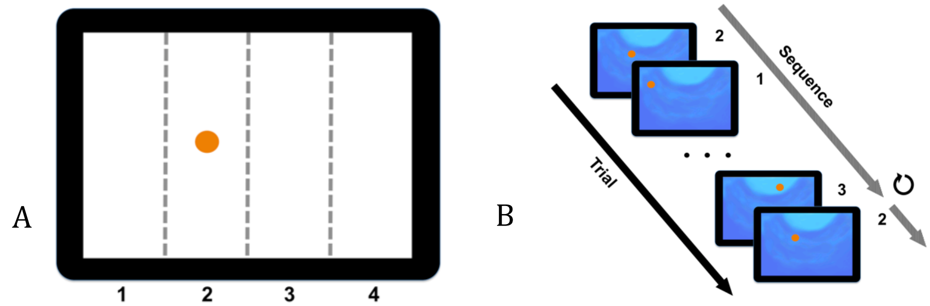


***Figure 4. Serial Response Time Task adapted for the present study.*** Participants will be presented with a bubble on each trial appearing in a random sequence (if in a random block) among four locations along the *x-axis* (as indicated in panel A). The bubble will always appear in a random *y-axis* position no matter the block type (as in panel B). During the learning blocks there is a specified cycling sequence that restarts from the start of the sequence (e.g., 2 in panel B) after the final item in the sequence (e.g., 3) has been popped.

The internal structure of our SRTT paradigm represented an additional adaptation from the typical paradigm with the variation of the task thought to more clearly reflect noisy and probabilistic stimulus distributions found in the natural environment (Destrebecqz and Cleeremans, 2001). The task was initially adapted from the aforementioned paradigmatic SRTT created by Nissen and Bullemer (1987). However, in contrast to many SRTT paradigms to date, we changed the relative vertical positions of the stimuli (in addition to the typical horizontal changes) making the task more probabilistic. Before describing the positions, first consider the concept and results of artificial grammar learning tasks introduced previously.

Within AGL paradigms, learners have been shown to transfer learning to a new set of stimuli as long as they adequately learn the rules governing the initial set (e.g., where recursions are possible or multiple paths could be taken). This provides for the possible generation of infinite sequences from infinite stimuli, but still having the same underlying rules governing position (grammar).

We adapted this convention to our SRTT paradigm, such that four spatial columns along the *x-axis* were considered to be the "grammatical positions" having the aforementioned sequence taken from Jimenez and Vazquez (2005) (see figure 4). The positions on the *y-axis*, on the other hand, were randomly generated by the software application. This variation in *y* represented the variability of items that could appear within each grammatical position. Thus, it is extremely unlikely that a participant will see the same sequence twice, nor would any other participant see the prior participant's sequences given the allowed variation in *y*. Thus, if reaction time to specific members of the sequence were to decrease over time, individuals would have indeed learned the underlying rules governing the horizontal position despite the dramatic variance in the vertical position.

In order to determine how the modifications we made influenced the classic demonstration of learning in serial reaction time tasks, a brief exploratory pilot study was conducted. We used this study to determine if learning effects were observed during our new and significantly modified serial reaction time task as described previously. However, analyzing the results of this pilot study was not as straightforward as analyzing results from the typical SRTT paradigm, as our task required significantly more movement and thus there were more potentially confounding variables to consider. While we of course could not consider all possibilities, we focus on a select few we suspected of potentially influencing the results, and the investigation of the descriptive data from the pilot study, we observed to impact the results.

During pilot testing, changes in distance covered had a substantial effect on reaction time, but not in the manner one might most clearly suspect. One might a priori suspect a linear slowing in reaction time across larger changes in magnitude. However, pilot participants actually showed very little variability in reaction time across magnitudes. This indicates that across larger distances participants moved much more quickly than they did to targets nearby. However, from a motor planning perspective, this makes intuitive sense; participants in our task had to move far more quickly to positions of higher magnitudes if they wanted to achieve optimal performance.

A second consideration for the effects of movement in our task is in reference to the probability of moving in one direction or the other hereafter referred to as *directional probability*. When playing whack-a-mole and the mole at the lowest left-hand side was just whacked, people will whack the next mole much more quickly if they prepare to move toward the area with the greatest number of probable moles. Similarly, in our task, we sought to determine if participants were preparing for movement to the next anticipated bubble based upon probable position of the next bubble. As no stimulus ever appeared in the same place twice in a row, participants were forced to prepare to move either to the left or to the right side of space. From the far left position and the far right positions, participants could only move one direction, and so we considered these as having an absolute probability of moving to the opposite side of space (i.e., $P = 1.00$). From the position to the left or right of center, participants were presented with far more movement probabilities. For example, from left of center they could move to the single remaining position on the left ($P$ = 0.33) or they could move to either of the two positions remaining to the right ($P$ = 0.66). If participants are directing their attention according to these rules, one could assume that they would move more quickly based upon these probabilities. Participants performing classic serial reaction time tasks may benefit from such knowledge as well but likely to a much smaller extent. Specifically, it may help participants to know that after pressing with their index finger, their only option is to next use one of their remaining three, but this likely won't influence performance as the button lies directly under each finger. In our task, participants have to move significantly larger distances to respond to stimuli. By introducing distance into the task, we provide for a more difficult situation from the classic task in which the participants will likely benefit to a much larger degree from knowledge of probability of movement direction.

It's important to note that certain levels of magnitude change are only possible from certain positions on the screen, and thus cannot be understood apart from a simultaneous consideration of the probability of movement direction. For the discussion, it will be helpful to consider the three levels of magnitude change as referring to three possible units of change on the screen from one of four positions. If the participant has just popped a bubble on the far left side of the screen, they can then assume with absolute certainty that they will be moving right, but the specific target is uncertain. From this far left position they may move one unit to the right (low magnitude change), two units to the right (moderate magnitude change), or three units to the right (large magnitude change). Then, the next bubble appears two units to the right. The participant's response to this bubble in the position to the right of center would thus constitute a moderate change in magnitude, and the right-ward direction they moved had an absolutely probability (1.00). From this new position to the right of center however, participants may move further still to the right, a one unit change with a directional probability of 0.33, or they may move one or two units to the left, each associated with a directional probability of 0.66. This is important, as movements of three-units (high magnitude) must always be associated with an absolute probability of directional movement, to one side or the other. Similarly moderate changes in movement may only be associated with a directional probability change of 0.66 or 1.00. In contrast, very small movements can be made from any position on the screen and can thus be associated with any of the three directional probabilities.

In order to determine the effects of these predictors on task performance, as well as how these disparate movement variables are related to our primary manipulation of sequence structure, we conducted an exploratory analysis of the pilot data.

### *4.1. Methods*

#### *4.1.1. Participants*

Seven young adult participants (3 male, 4 female; ages 21 – 28) participated in the brief pilot study. The participants were graduate student volunteers and acquaintances of the experimenter or one of the experimenter's colleagues. Prior to participation in the task, the volunteers were naïve to the study manipulation and had not participated in a serial reaction time task in the past.

*4.1.2. Apparatus*

A Microsoft Surface Pro® tablet computer in landscape orientation was intended for use the presentation of our SRTT but was not available at the time in which the pilot study was taking place. An available conventional touchscreen monitor connected to a Windows XP computer was thus used instead to display the task. The computer's monitor was a 3M™ 17" LCD touchscreen that was tilted at an approximately 30-degree angle to allow the participants greater ease of use. The screen's resolution was approximately 96.42 pixels per inch.

*4.1.3. Stimuli*

The SRTT consisted of four blocks of 72 trials (for a total of 288) with a single bubble (120 pixel diameter, ~31.61mm) appearing on each trial. There was an initial practice block of 36 trials at the beginning of the experiment to allow the participants to familiarize themselves with the testing environment. The screen was divided into four equally spaced *columns*, with bubbles appearing in the center of the columns (see figure 3), resulting in the bubbles being centered on the horizontal axis at +/- 126.45mm or +/- 42.15mm. During each trial of the experiment, the bubble positions on the *y-axis* were randomized such that the bubble remained in the center of the vertical axis, or was offset from center by +/-20.68mm. Upon “popping” the bubbles, the participant was given an auditory reward (sound of a bubble popping).

*4.1.4. Procedure*

Participants were first given a brief demonstration on how to pop the first three bubbles , and were shown what would happen as a result (i.e., the visual and auditory reward). Participants were then allowed to interact with the task for the remainder of the trials without further instruction. After each of the first three blocks, a “continue” button would appear on the screen. Participants were given the option to take a short break of 15-30 seconds when the continue button was displayed on the screen. The participants were asked to press continue as soon as they were ready to proceed. This procedure was followed for the remainder of the four blocks. After the fourth block pressing the continue button exited the SRTT experiment. The entire task took approximately 4-minutes to complete.

Blocks 1 and 4 had randomly distributed bubbles along the *Columns* as well as the *Rows* with the constraint that the horizontal position could not be repeated twice in a row. The second and third blocks constituted the learning blocks and contained 8 repetitions of a 12-item sequence. Specifically, the sequence proceeded through the columns in the following manner: 2-1-4-1-3-4-2-3-1-2-4-3. The starting point in this sequence was randomized across participants. Once the sequence ended it would start from where it began until all six iterations of the sequence were complete (see figure 4 for an example).

*4.2. Statistical Approach*

We utilized a mixed-effects model building approach in statistical analysis of the results. Mixed-effects models allow one to analyze changes in task performance while simultaneously capturing multiple significant grouping factors that may impact data analysis and subsequent model fit. As compared to typical linear modeling and ANOVA approaches, mixed-effects modeling allow for inclusion of crossed-random effects in the same analysis (e.g., effect of participant as well as item effects) (Bates, 2010). Additionally, mixed-effects models remain robust in the face of missing values and sparce matrices of data. Both of these considerations were particularly relevant to our understanding of performance on our version of the SRT task given potentially influential task specific effects which may influence our results, but with mixed-effects modeling, can be appropriately accounted for. The R statistical language (version 3.0.2) was utilized for all analyses (R Core Team, 2013) in the present study and the "lme4" package (version 1.0-5 ) within R (Bates, 2013) was utilized to develop mixed-effects models, developed using the full data set. The "lmer" function of the lme4 package was used to fit the general linear mixed models using random estimation maximum likelihood (REML) in order to first inspect the models for multi-colinearity and improper specification of the effects. Maximum likelihood models describe the probability of the data given a specific model and parameter values. Upon successful inspection, appropriately specified models were then recomputed using maximum likelihood estimation (MLE) for comparison to other models. This procedure was followed as model comparisons between REML models leads to less certainty that a comparison of model fit is accurate, though MLE is thought to excel in this case. On the other hand inspection and specification of the specific parameters to be interpreted from the model are argued to be more stable when specified via REML. Thus, any model comparison is computed under MLE, but all parameter estimates are reported from REML models

In order to specify the unconditional models, including only random effects, we undertook a hierarchical model building approach in which we started with a base model including only the intercept and the random effect of participant on the intercept, then adding and eliminating random effects in order of theoretical importance, before modeling theoretically important fixed effects (Bates, 2010). To determine if each random effect significantly impacted the model's fit of the data, we compared hierarchical models utilizing the Likelihood Ratio Test (LRT) with separate ANOVAs. The LRT allows for a comparison of the goodness of fit between two models when one or more effects have been included or removed, with goodness of fit estimated using a chi-squared distribution (Mirman, 2014).

Each unconditional model was compared against the following base model: "*Speed ~ 1 + (1|participant) + (1|trial)*", which can be read as, "Speed is a function of the intercept, allowing for change in the intercept by participant and by trial. "

Upon finding the appropriate model specification, parameters of the model(s) are reported with parameter-specific p-values, estimated using the normal approximation (i.e, z-values) provided by the "multcomps" package (version 1.3-1) (Hothorn, Bretz, & Westfall, 2008) in R. The multcomps package additionally allows for automatic significance value correction utilizing the correction method of choice.

Planned and post-hoc comparisons of fixed effects were made utilizing family-wise Bonferroni correction.

It is of note that whenever the response variables for a model were based on reaction time, the values were transformed to a log scale to reflect the positively skewed and non-scalar nature of reaction time data (Ratcliff & Murdock, 1976). Estimates of model specific parameters should therefore be interpreted with caution as they represent changes to the logged values of participant speed. It should also be noted that on this log scale, values will seem far smaller than one would anticipate for estimates.

For simplicity sake, one can assume that whenever we tested a fixed effect or an interaction, per-participant variability was allowed for that effect and modeled as such. As each level of the fixed effects occurred within subject, the variability structure was modeled with this in mind (Mirman, 2014). To jump ahead and use *block* in the model as an example, per-participant variability was modeled viz-a-viz "block + (1|participant) + (1|participant:block)". This using the lmer package allows variability for the participant in the intercept without the slope, as well as a per-participant interaction at the intercept by block, thus taking into the participant-by-block interaction. It should be noted that for more complex models involving interactions of the fixed effect, per participant variability was allowed in all lower terms as well as for the interaction term. For example: "1 + block*probability + (1|participant) + (1|participant:block) + (1|participant:probability) + (1|participant:block:probability)" .

For ease of reference we report the equations submitted for the model building approach at the end of each relevant section.

*4.2.1. Random effects compared for models of SRT task performance.*

Logged speed to targets in milliseconds per millimeter to targets ($ms/mm_{Log}$) was considered our primary experiment-wide dependent variable. The decision to conduct analyses on speed to target rather than absolute reaction time is based upon the differential speed of performance that is not necessarily reflected in absolute measures of reaction time. The following variables are included at various stages in the model building process. *Participant* represents intra-individual variability. *Trial* representing potential fatigue effects associated with performance of the task over time with the trial variable representing trials from 1 – 280. The participant and trial level variability, modeled as random effects, were always included in the model and were never subject for removal as result of hierarchical comparisons of the random effects. *Jitter* represents the random movement among three possible positions on the vertical-axis. *Sequence repetition* specifies how many times the sequence had been repeated (1 – 6). For the blocks 1 and 4 without a consistent sequential pattern, the sequence repetition variable simply collapsed sets of 12 trials into 6 separate bins, indicating speeding up or slowing down that may have occurred within a block. The *Left vs. Right* binomial factor indicated left versus right movement trajectory. *Magnitude* indicates the absolute change in magnitude distance required from moving to the previous position to the current. *Directional probability,* referred to for simplicity as *probability*, indicates the probability of movement direction that was available when moving from the previous bubble position to the current (0.33, 0.66, 1.00).

#### *4.2.2. Fixed effects included in the models of SRT task performance.*

*Block* is a five level factor representing, in order, an initial practice block of trials (practice), a random block containing no repeating stimulus pattern (block 1), a *learning* block containing the repeating pattern of horizontal positions with random vertical jitter (block 2), a second learning block (block 3), and a final random block (block 4). For all statistical analyses, the practice block was omitted from the dataset. Block 1 is specified as the base contrast for the multinomial predictor block. *Directional Probability* also referred to in the text as *probability* or shortened to *P*(X), where X represents the specific probability. Probability is a three-level multinomial factor with levels 0.33, 0.66, and 1.00, with *P*(0.33) serving as the base contrast.

### *4.3. Results*

#### *4.3.1. Data cleaning.*

Prior to conducting the modeling and interpretation of these data, reaction times were capped based upon data points falling 2.5 standard deviations above the median reaction time for each participant. The median reaction time was used as it represents a “robust estimator” meaning it is less sensitive to skew in the distribution of response latencies due to noise in the participants’ responses

(Ratcliff, 1993). Values falling above the median reaction time were then capped at 2.5 SD above the participant's median. The choice to utilize data truncation was based upon Ratcliff's recommendation to minimize the affects of outliers on data analysis (Ratcliff, 1993). We argue 2.5 standard deviations above the mean of each participant's data is far enough outside the average response range that it minimizes both the effect of extreme outliers, as well as the likelihood of removing responses representing the construct under investigation (Ulrich & Miller, 1994) (i.e., implicit learning observed via a shortening of response times). This technique resulted in few trials requiring capping (Mean = 1.45% (0.05%), Range = [0.7%, 2.19%]).

The first trial of each block was not included in the analysis as due to a computer error, the output of the axis positions for the first trial was incorrect. Upon further investigation, it was found that the first trial of the blocks including the repeating sequential were visually confirmed to be in the appropriate location; thus it was only the output of the first bubble's position that was not accurately reported in the data output.

*4.3.2. Modeling the random effects*

Through the analysis we found jitter on the vertical axis did not significantly affect performance above the base model (BASE vs. M1: $\chi^2$ (1) = 0, p = 1.00). Left vs. right-ward movement (irrespective of directional probability) also did not significantly contribute to the model (BASE vs. M2: $\chi^2$ (1) = 0.01, p =0.94), nor did sequence repetition (BASE vs. M3: $\chi^2$ (1) = 0, p = 1.00). Thus, the model compared against for the initial fixed effects, was the base model (BASE).

BASE: $ms/mm_{log} \sim 1 + (1|participant) + (1|trial)$

M1: BASE + $(1|jitter)$

M2: BASE + $(1|$l*eft vs. right*$)$

M3: BASE + $(1|sequence)$

*4.3.2. Modeling the fixed effects*

When including magnitude in the model as a fixed effect, we found that magnitude plays a significant role in the task and thus significantly improved the model fit of the data (BASE vs. M4: $\chi^2$ (3) = 1559, p < 0.001). We then modeled the effect of *block* which above magnitude, further improved model fit (M4 vs. M5: $\chi^2$ (4) = 60, p < 0.001). There was no effect of allowing for an interaction between block and magnitude (M5 vs. M6: $\chi^2$ (5) = 6.25, p = 0.28), indicating both can be interpreted separately as main effects. The estimates for the analysis including magnitude and block as fixed effects are represented by table 2.

**Table 2. Parameter Estimates for the Fixed Effects Model of Magnitude and Block in the Pilot Study**

| | Estimat | Std. | *z* | *p* |
|---|---|---|---|---|
| Intercept | 0.681 | 0.048 | 14.17 | 0.000 |
| Magnitude: Low vs. Moderate | -0.296 | 0.016 | -18.03 | 0.000 |
| Magnitude: Low vs. High | -0.400 | 0.017 | -23.09 | 0.000 |
| Block: 1 vs. 2 | -0.039 | 0.019 | -2.00 | 0.045 |
| Block: 1 vs. 3 | -0.058 | 0.019 | -2.97 | 0.003 |
| Block: 1 vs. 4 | -0.034 | 0.019 | -1.76 | 0.078 |

We then analyzed the influence of directional probability and its influence on block effects. For this analysis we included magnitude as a random effect so as to allow for an understanding of the association between probability and block on performance apart from magnitude change (M7). Probability improved model fit beyond the base model, including the random effect of magnitude (M7 vs. M8: $\chi^2$ (3) = 284, p < 0.001). Beyond probability, inclusion of block further improved model fit (M8 vs. M9: $\chi^2$ (4) = 65.7, p < 0.001).

**Table 3. Parameter Estimates of Logged Speed for the Fixed Effects Model of Block and Probability in the Pilot Study**

| | Estimates | Std. Error | *t* |
|---|---|---|---|
| Intercept | 0.346 | 0.149 | 2.32 |
| Block 1 vs. 2 | 0.030 | 0.029 | 1.03 |
| Block 1 vs. 3 | 0.037 | 0.029 | 1.26 |
| Block 1 vs. 4 | 0.016 | 0.029 | 0.54 |
| Probability 0.33 vs. 0.66 | 0.067 | 0.030 | 2.24 |
| Probability 0.33 vs. 1.00 | 0.139 | 0.029 | 4.77 |
| B1 vs. B2 for P0.33 vs. P0.66 | -0.124 | 0.034 | -3.66 |
| B1 vs. B3 for P0.33 vs. P0.66 | -0.167 | 0.034 | -4.92 |
| B1 vs. B4 for P0.33 vs. P0.66 | -0.086 | 0.035 | -2.49 |
| B1 vs. B2 for P0.33 vs. P1.00 | -0.058 | 0.032 | -1.80 |
| B1 vs. B3 for P0.33 vs. P1.00 | -0.078 | 0.034 | -2.40 |
| B1 vs. B4 for P0.33 vs. P1.00 | -0.044 | 0.033 | -1.34 |

Finally, an inclusion of the interaction between both probability and block again improved model fit beyond the model indicating only main effects for each (M9 vs. M10: $\chi^2$ (7) = 49, p < 0.001). The fixed effects table for this final model (M10) are presented in table 3.

M4: BASE + *Magnitude* + (1|*participant:Magnitude*)

M5: M4 + *Block* + (1|*participant:Block*)

M6: M5 + *Magnitude:Block* + (1|*participant:Magnitude:Block*)

M7: BASE + (1|*magnitude*)

M8: M7 + *Probability* + (1|*participant:Probability*)

M9: M8 + *Block* + (1|*participant:Block*)

M10: M9 + *Probability:Block* + (1|*participant:Probability:Block*)

*4.4. Discussion*

The classic serial reaction time task requires only a button press and for participants to rest their hand in a single position over the course of the experiment whereas ours requires substantially more movement. As such, we explored the influence of several variables that, under typical SRT task conditions, need not be considered. Specifically, we investigated the influence of magnitude of movement required from trial to trial, the effects of left-ward vs. right-ward movement, as well as the probability of movement to one direction rather than the other. In the classic task, magnitude of movement is not necessarily relevant as participants rest their fingers on the keys required to respond to stimuli. In contrast, to respond in our task, participants are required to cover various distances from trial to trial given the variation along the horizontal axis as well as the random jitter on the vertical axis.

The above movement conditions and their effects on our results are depicted clearly in the top left and right panels of figure 5 (pilot results on the left, and jumping far ahead, primary experimental results on the right). Figure 5 indicates that collapsing performance across all magnitudes in our task, or failing to consider magnitude in modeling performance, is ill advised, as there is a clear difference in participant speed to different magnitudes. However, very slight magnitude change, associated with minimal jitter on the vertical axis, didn't significantly impact performance, but relatively larger magnitudes, those associated with a change on the horizontal axis, did. Indeed, nearly equivocal performance across the positions of the vertical axis was observed for the pilot study as well as during the experiment proper (see both the top pilot and primary experiment panels of figure 5). For the sake of convenience of discussion and interpretation of results we refer to a scale of relative changes in magnitude collapsing across the very small variability associated with change on the vertical axis. Magnitude changes between 84mm and 94mm are referred to as *low*, changes between 168 and 174mm are considered *moderate*, and changes between 252 and 257mm are considered *high*.

The results as well as graphical depiction of these pilot data, allowing for an interaction between block and probability, indicate that indeed participants in the pilot study revealed a classic effect of learning, but only for directional probability of 1.00 and for 0.66 (left panel of figure 6 – note the panel on the right will be explained in due time during the results and discussion of the primary experiment). Specifically, they showed increased speed in performance from block 1 to block 3 after the task environment provided structure to the variability across trials (i.e., a consistent horizontal sequence began during block 2 and continued through block 3). However, when this structure in the environment was removed and sequence order was returned to random (as going from block 3 to 4), participants showed a performance decrement for probability 0.66 and 1.00. Of these two probabilities, the sequence was most helpful for a directional probability shift of 0.66. In contrast, directional probability of 0.33 did not appear significantly impacted by the predictability in the environment as much as the higher order probabilities. Together the results of these pilot data indicate that under some conditions of directional uncertainty, there is much greater room for benefit. At lower probabilities the influence of learning appears to have a greater effect than at the higher level (0.66 vs. 1.00), but with very high levels of uncertainty (0.33), less benefit is provided.

Given the exploratory findings in the analysis of the influence of these movement related factors in the pilot data, we concluded that both magnitude and probability should be included as factors in our primary experiment, and we thus conducted the analyses and interpreted the findings in our primary experiment in light of these effects.

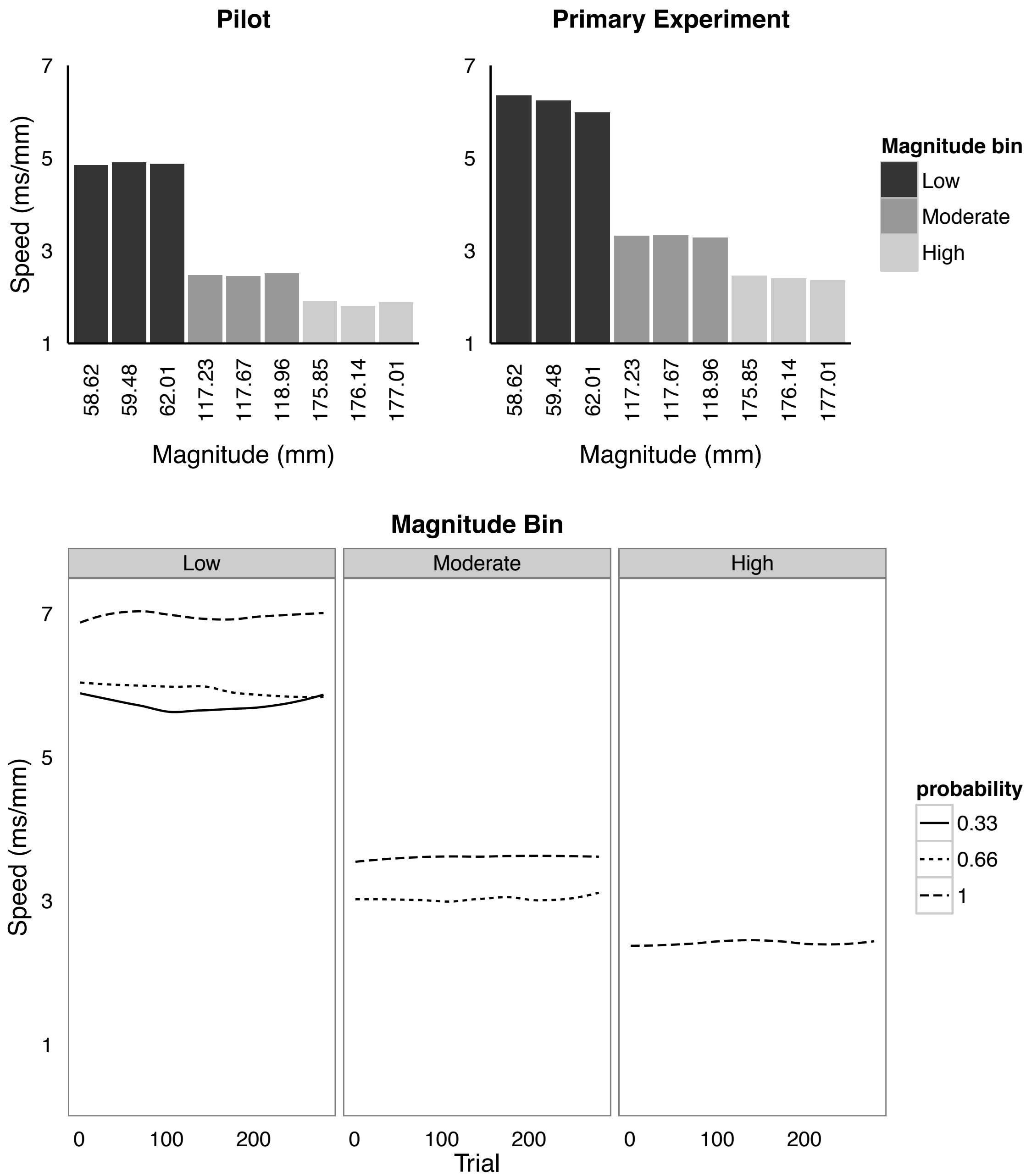


*Figure 5.* ***Results of the Pilot and Primary Experiments Reflecting Movement Related Change in Performance.*** Results from the pilot (top left panel) and primary (top right and lower panels) experiments indicate that magnitude of movement significantly contributed to performance differences such that participants moved far faster when the bubbles to be popped were further away than the ones nearby. The lower panel from the primary experiment reflects changes in magnitude as necessarily dependent upon directional probability (probability of moving to one side of space over the other on a particular trial) given our task constraints.

## Chapter 5. Primary Experiment Part 1: Extension of the Results to a Tablet Version and Isolation of Individual Differences

As our new version of the serial reaction time task indeed revealed a stereotypical learning effect in the pilot study under certain conditions of movement, it may well be a successful candidate task in addressing our primary goals in conducting this study. That is, we may be able to determine if individual differences in implicit learning exist (hypothesis 1), and determine if these individual differences are meaningfully associated with higher order cognitive abilities (hypothesis 2). As indicated in the introduction, findings supporting either hypothesis will be important in further determining the utility of implicit learning as a construct worth clinical assessment in a cognitive battery.

We therefore proceeded with the primary experiment utilizing the serial reaction time task on a tablet version of the task as originally planned and attempt to extract metrics of individual differences. Findings of individual differences was a requirement in order to proceed to hypothesis 2, as without measurable and somewhat stable individual differences, any correlation with other higher order tasks could not be determined.

Stepping back further, as our primary goal is to adapt a version of the SRT for use with a touchscreen tablet, hypothesis 1 could not be addressed without first replicating the study using a tablet computer rather than a conventional touchscreen. We move away from conventional touchscreens, as these devices are not available in many clinical locations and are mostly absent from the bedside in clinical settings. Additionally Tablet computers are cheaper and far more portable, and are thus more likely to lend well to the clinical setting.

However, the pilot study was quite disparate from the current in terms of presentation of the stimuli. This is important, as although the response mechanisms were intended to be similar no matter the size of the touchscreen, the screen size of the conventional touchscreen used for the pilot study was far bigger than that of the average touchscreen. We attempted to maintain the same aspect ratio for the stimuli, thus for the smaller screen the stimuli were smaller, spaced more closely together, and required greater precision in response than on the conventional touchscreen. This was done as we desired a greater level of precision in responding for the participants, as a direct. translation of the bubble size from the conventional touchscreen would have been too big

Given these differences, results based upon the pilot study cannot be considered equivocal to the primary experiment, as participants were required to move a far greater distance from target to target during the pilot. Nevertheless an analysis of the potential similarities between the two tasks was deemed informative, particularly given the surprising influence of magnitude and directional probability seen during the pilot study.

In order to determine whether or not our replication of the task on the tablet continued to measure learning, retesting of the entire statistical modeling procedure from the pilot study was undertaken, assuming no replication. However, rather than restest magnitude and probability as fixed effects, we do test these variables as random effects, comparing them against the baseline model including only participant and trial effects. Upon finding a learning effect however, we then sought to extract the individual differences.

As rate of learning in typical SRT paradigms is typically followed by a decrement to performance when the learned pattern is removed, we also investigated the difference between block 3 and block 4, assuming that if participants showed a learning effect, they would be significantly slower at block 4 as compared to block 3. This is in essence, the classic "V" shape pattern shown in typical SRT paradigms, indicating continuous learning on the task, but then a significant performance decrement when the learning sequence is removed.

To determine if the learning rate differed from one probability to the other, we compared block 1 to block 3 at each level of probability. As the single most common measure of learning reported in SRT tasks is the effect of removal of the sequence, we focused the remainder of our analyses on the differences between block 3 and block 4 for each probability. Specifically, we not only compared block 3 to block 4 at each level of probability (three contrasts), we also computed differences in slope of this change between blocks 3 and 4 for probabilities 0.33 vs. 0.66 as well as for 0.33 vs. 1.00. To interpret these effects with greater clarity we additionally compared probabilities at baseline level of performance (block 1) to determine how baseline performance was influenced by probability. This resulted in a total of 10 planned contrasts, compared against a Bonferroni corrected *p-value* of 0.005.

In order to measure individual differences in our task, it can be observed that simply subtracting performance on block 4 from performance on block 3 removes the subtly with which learning is demonstrated in the context of directional probability. Therefore we sought to take these important variables into account when specifying individual differences.

To do this we rely upon the classic definition of learning on SRT tasks. Specifically, we argue that individuals who demonstrate "learning" in our task should improve significantly from block 1 to block 3 and they do so to a far larger degree than individuals with lower rates of learning. Our argument additionally specifies that for individuals who demonstrated better learning from block 1 to block 3, they should additionally show significantly greater decrement transitioning from block 3 to block 4. This is because if participants learned the sequence to a relatively larger degree than the average participant, they should similarly have more to lose when that "well-learned" sequence is removed (as is the case for block 4). Thus, the individual's learning rate should be positively correlated with their performance decrement as compared to average. As probability is shown to have a differential role on learning effect, we correlate these components in order to compare learning effects to decrements across each probability.

Contrasts between the final model parameters, comparing the influence of probability across blocks, were planned based upon the classical pattern of performance observed for the Serial Reaction Time Task and kept to the minimum possible in order to understand the interaction. Specifically, as we were interested in learning rate, we investigated the improvement in speed that occurred from blocks 1 to 3. In order to more clearly understand the effects, Bonferroni corrected post-hoc multiple comparisons were conducted to explain the interaction.

*5.1. Methods*

*5.1.2. Participants*

Forty-five young adult participants were recruited from Drexel University. Participants were recruited through either an online advertisement system or flyers posted throughout Drexel's campuses. The online advertisement system is made available to students in psychology courses seeking to gain extra-credit for a course. All methodology used in the present study was approved by the Internal Review Board at Drexel University. It is of note that four participants were excluded from the study; two participants left during the task due to conflict in schedule and inability to return for follow-up testing and two participants performed the task while being substantially distracted (self-reported) by nearby construction throughout the experiment session.

*5.1.2. Apparatus*

A Microsoft Surface Pro® tablet computer in landscape orientation was used in the presentation of our SRTT. The Surface Pro® utilized was a first generation (February, 2013) 64gb model, having a refresh rate of 60Hz, 4GB of RAM, and an Intel Dual-core 1.7 GHz processor.

*5.1.3. Stimuli*

The stimuli and 5-block design of the SRTT (practice, block 1 – block 4) were identical in every respect to the design used for the pilot save for the size of the bubble and it's spacing about the screen. For each of the 288 trials, a single bubble (120 pixel diameter, ~14.65mm) appeared on each trial. The Surface® screen was divided into four equally spaced *Columns*, with bubbles appearing in the center of the *Columns* (see figure 3), resulting in the bubbles being centered at ~87.92mm (720 pixels) or ~29.31mm (240 pixels) to the left or right of the center of the screen's horizontal axis. During each trial of the experiment, the bubble positions on the *y-axis* (hereafter *Rows*) were randomized such that the bubble remained in the center of the vertical axis, or was offset from center by +/- 10.11 mm (82.8 pixels up or down).

### *5.2. Statistical Approach*

We utilized an identical mixed-effects modeling approach as presented in chapter 4 for the pilot study, with the same definitions of variables included in both the fixed and random effects. Further, the same model building approach was utilized, and the same techniques for assessing the model fit in R were used (Bates, 2010; Mirman, 2013).

In addition to the aforementioned random and fixed effects included in the model, Bonferroni corrected post-hoc multiple comparisons were conducted to explain the interaction between probability and block. The alpha-level for the contrasts was thus set to 0.005, as 10 comparisons were conducted.

Individual differences in the implicit learning task were extracted as individual "learning rates" representing each individual's performance during block 3 subtracted from their performance during block 1. Performance decrement based upon removal of the predictable stimulus patterns was extracted by subtracting each individual's performance during 4 from their performance during block 3. As these two components should be correlated for learners (high learning rate followed by a high performance decrement), a simple regression was conducted within R using the base R package, regressing each learning rate component on each decrement rate component for each probability, creating a correlation matri.x. For this correlation matrix covariates were not considered.

## *5.2. Results*

### *5.2.1. Data cleaning.*

Prior to conducting the modeling and interpretation of these data, reaction times were cleaned in the same manner as in the pilot study and then transformed to a log scale for further processing. Few trials required capping (Mean = 1.45% (0.05%), Range = [0.7%, 2.19%]). Similarly to the pilot study, the first trial of each block was removed due to computer error.

### *5.2.2. Modeling the random effects*

Model building indicated that no improvement in goodness of fit was provided by including the effect of specific vertical position (jitter) (BASE vs. M1: $\chi^2$ (1) = 0.00, p = 1.00), left vs. right directional transition on the intercept (BASE vs. M2: $\chi^2$ (1) = 0.14, p = 0.71), nor sequence repetition (BASE vs. M3: $\chi^2$ (1) = 1.77, p = 0.18). As in the pilot study, a significant influence of the random effect of magnitude was found

(BASE vs. M4: $\chi^2$ (1) = 16199, p < 0.001), as was the random effect of probability (M4 vs. M5: $\chi^2$ (1) = 1617, p < 0.001).

BASE: $ms/mm_{log} \sim 1 + (1|participant) + (1|trial)$

M1: BASE + (1|*jitter*)

M2: BASE + (1|l*eft vs. right*)

M3: BASE + (1|*sequence*)

M4: BASE + (1|*magnitude*)

M5: M4 + (1|*probability*)

*5.2.3. Modeling of the fixed effects of block*

As in the pilot study, we found a significant influence of block beyond the effects of the random effects in the model ($\chi^2$ (4) = 54.9, p < 0.001). The fixed effects table for the model of block and the random effects of participant, trial, magnitude, and directional probability are indicated in table 4.

**Table 4. Parameter Estimates of Logged Speed for the Fixed Effects Model of Block in the Primary Study**

| | Estimat | Std. | *t* |
|---|---|---|---|
| Intercept | 0.531 | 0.138 | 3.84 |
| Block 1 vs. 2 | -0.002 | 0.004 | -0.57 |
| Block 1 vs. 3 | -0.003 | 0.004 | -0.93 |
| Block 1 vs. 4 | -0.006 | 0.004 | -1.61 |

*It should be noted that Std. Error is rounded to the nearest thousandth, and so identical values as reported in the table should not be taken to mean that the errors are in fact equivocal in the actual data.

Though the model of block improved model fit, the estimates provided for the effect of block when interpreted apart from the effects of probability as in the pilot study, do not provide a very accurate description of the pattern observed when visualizing these data. However, the differential effects of block were clearly demonstrated upon brief inspection of the estimates for the random effects generated for each probability by the aforementioned model (e.g., when extracting the estimates from the model for each participant at each block, for each probability, collapsing across magnitude). In order to more clearly understand these results, we thus conducted a similar analysis to the one conducted for the pilot study, including probability as a fixed rather than random effect.

*5.2.4. Modeling of the interaction between block and probability*

Similar to the pilot study, the influence of block added improved model fit beyond probability and the random effects (participant, trial, magnitude), ($\chi^2$ (4) = 57.5, p < 0.001). As a second replication of the pilot results, the influence of the block by probability interaction further improved model fit beyond either included as only main effects ($\chi^2$ (7) = 26.7, p < 0.001). The fixed effects for the final model are provided in table 5 and the graphical depiction is provided by the right panel of figure 6.

The results of the post-hoc comparisons indicate that performance during the first block is not significantly different between probability 0.33 and 0.66 (*Estimate* = 0.005 *SE* = 0.008, *p* = 0.429), but is significantly different between probability 0.66 and 1.00 (*Estimate* = 0.068 *SE* = 0.005, *p* < 0.001). Specifically, speed to stimuli associated with an absolute movement trajectory (1.00) is far slower than to either probability 0.66, or 0.33. As the estimate for probability 0.33 indicates it is slightly faster than for probability 0.66 during block 1, it is assumed that speed for probability 0.33 at block 1 is also faster than for probability 1.00, though this comparison was not made. The results additionally indicate a differential effect of learning trials (block 1 vs. block 3) on performance for probability 0.33 (*Estimate* = -0.021 *SE* = 0.006, $p < 0.001$), but not for 0.66 (*Estimate* = -0.002 *SE* = 0.005, $p = 0.700$) or 1.00 (*Estimate* = -0.002, *SE* = 0.004, $p = 0.62$). Specifically, the rate of learning (i.e, increased performance from block 1 to 3) is significant for probability 0.33, but not for either probability 0.66 or 1.00.

**Table 5. Parameter Estimates of Logged Speed for the Fixed Effects Model of Block and Probability in the Primary Study**

| | Estimates | Std. Error* | *t* |
|---|---|---|---|
| Intercept | 0.507 | 0.136 | 3.74 |
| Block 1 vs. 2 | 0.030 | 0.006 | -2.96 |
| Block 1 vs. 3 | 0.037 | 0.006 | -3.46 |
| Block 1 vs. 4 | 0.016 | 0.006 | -1.30 |
| Probability 0.33 vs. 0.66 | 0.067 | 0.006 | 0.79 |
| Probability 0.33 vs. 1.00 | 0.139 | 0.005 | 13.14 |
| B1 vs. B2 for P0.33 vs. P0.66 | -0.124 | 0.007 | 2.14 |
| B1 vs. B3 for P0.33 vs. P0.66 | -0.167 | 0.007 | 2.82 |
| B1 vs. B4 for P0.33 vs. P0.66 | -0.086 | 0.007 | -0.58 |
| B1 vs. B2 for P0.33 vs. P1.00 | -0.058 | 0.007 | 3.53 |
| B1 vs. B3 for P0.33 vs. P1.00 | -0.078 | 0.007 | 3.62 |
| B1 vs. B4 for P0.33 vs. P1.00 | -0.044 | 0.007 | 1.21 |

*It should be noted that standard error is rounded to the nearest thousandth, and so identical values as reported in the table should not be taken to mean that the errors are in fact equivocal in the actual data.

The performance decrement observed with removal of the sequence (i.e., reduction in speed from block 3 to 4) was observed for only probability 0.33 (*Estimate* = 0.013, *SE* = 0.006, $p = 0.034$), while probability 0.66 actually showed increase in performance from block 3 to 4 (*Estimate* = -0.01, *SE* = 0.005, $p = 0.033$), and at probability 1.00 showing reasonably stable performance from block 3 to 4 (*Estimate* = -0.002, *SE* = 0.004, $p = 0.60$). However, it should be noted that the decrease in performance for probability 0.33 from block 3 to 4, as well as the increase in performance for probability 0.66 was not significant after Bonferroni correction (i.e., $p < 0.005$). In contrast, the difference in slope between block 3 and 4 for probability 0.33 as compared to 0.66 was significant even after Bonferroni correction (*Estimate* = -0.024, *SE* = 0.007, $p < 0.001$), supporting the contention that performance decreased for 0.33 from block 3 to 4, but increased for 0.66 across the same interval. The performance decrement from block 3 to 4 was also different between probabilities 0.33 and 1.00 but not significantly so after Bonferroni correction (*Estimate* = -0.016, *SE* = 0.006, $p = 0.018$).

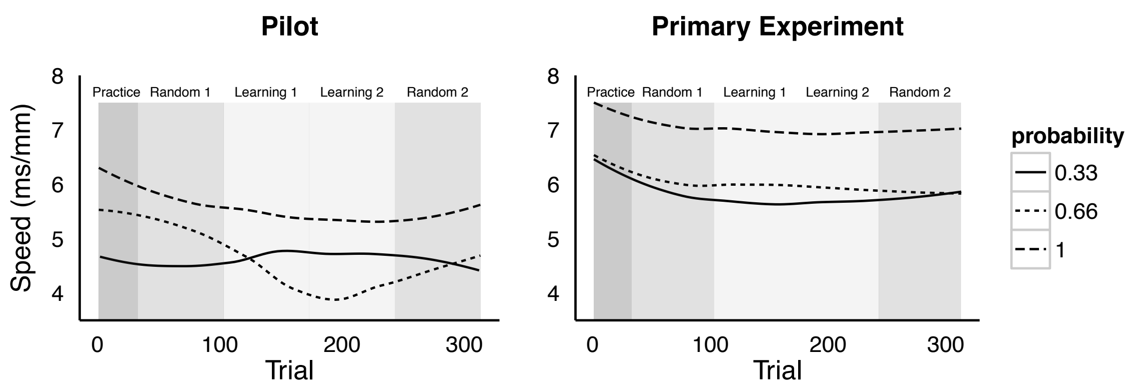


***Figure 6. Results from the pilot and primary experiments across directional probabilities.*** Data from both the pilot data (left) and experiment data (right) are shown, with speed as a function of the interaction between probability and block. The graphs demonstrate the significant learning effects for some probabilities (0.66 and 1.00 in the pilot experiment; 0.33 and 1.00 in the primary experiment). The graphs also demonstrate the consistently lower performance across experiments with higher directional probability, demonstrated even through the practice trials. It is of note that these panels represent only the lowest magnitudes in each experiment, as given task constraints, all probabilities cannot be compared for each magnitude.

Correlations between the learning rates and performance decrements for each probability are shown in the correlation matrix in table 6. This correlation matrix indicates that for the probabilities that demonstrated significant learning effects upon comparison of planned contrasts of the fixed effects model (0.33 and 1.00), there is a significant positive correlation between each rate of learning (block 1 minus block 3) and performance decrement (block 4 minus block 3) component. It is important to remember that no performance decrement was observed at probability 1.00 for the group as a whole. However, the correlation between the learning rate of 1.00 and the performance decrement for 1.00 indicates that some individuals likely did demonstrate learning during this probability even if the group did not. This contention is further supported by the large correlation between learning effect for probability 0.33 (the only probability that did demonstrate learning across the group) and both the rate of learning and the performance decrement observed for probability 1.00. Visual inspection of the residuals vs. the predicted values of these comparisons indicates that the predictors in each linear regression are well specified without demonstrating significant multi-colinearity or inappropriately modeled components.

**Table 6. Correlation between improvement from blocks 1 and 3 with decrement from block 3 to 4 compared by probability**

| | *P*(0.33): B1-B3 | *P*(0..33): B3-B4 | *P*(0.66): B1-B3 | *P*(0.66): B3-B4 | *P*(1.00): B1-B3 | *P*(1.00): B3-B4 |
|---|---|---|---|---|---|---|
| *P*(0.33): B1-B3 | . | . | . | . | . | . |
| *P*(0.33): B3-B4 | -0.830** | . | . | . | . | . |
| *P*(0.66): B1-B3 | -0.315* | -0.142 | . | . | . | . |
| *P*(0.66): B3-B4 | -0.080 | -0.137 | -0.055 | . | . | . |
| *P*(1.00): B1-B3 | -0.584*** | -0.571* | -0.136 | 0.009 | . | . |
| *P*(1.00): B3-B4 | -0.377* | -0.505* | -0.287 | -0.290* | -0.396** | . |

Directional probabilities are presented in the table as *P*(XX) where XX is the particular probability under investigation. B1-B3 indicates the value of subtracting participants' reaction time from the third block from the first block, with positive values representing higher rates of learning. B3-B4 indicates the value of subtracting participant reaction time from of the third block from the fourth block with higher values representing higher performance decrement. Asterisks represent levels of significance for correlations, reported as * = $p < 0.05$, * = p < 0.01, and *** as $p < 0.001$.

### *5.3. Discussion*

It has thus far been demonstrated that an individual's rate of learning is significantly correlated with their performance decrement. However, this is only observed for probability 0.33 and 1.00, corroborating the indicated significant contrasts performed on the final model. We may now therefore reasonably interpret individual differences in these components as reflecting individual differences in implicit learning. Further, it is also clear that we may utilize the single greatest predictor of learning in our task to predict performance on these other tasks. Specifically, as the rate of learning for probability 0.33 predicts all other observed learning components, we choose this component as our primary measure of learning in our serial reaction time task. It is from this context that we compare performance on the implicit learning task to other measures of higher-order cognition.

## Chapter 6. Primary Experiment Part 2: Are The Individual Differences Observed in Our Implicit Learning Task Meaningful?

Upon extracting metrics of individual differences we now move to part 2 of the exploration of individual differences in our serial reaction time paradigm. During part 2, we attempt to extract a measure of individual differences in a common linguistic ability for our participants. As there are not any normative tasks available for adults that analyze the components of language processing using a significantly fine-grained analysis, we compare our task to one of those available in the research literature. As we rely upon a research task that has no normative measure of individual differences, we must again ensure the task measures what it is we intend it to measure. In this vein, we decided to utilize the Syntactic Comprehension created by Misyak and Christiansen (2007). We use this task, as in previous work by the authors, the task has been found to be associated with implicit learning ability. We thus sought to replicate their findings using our new implicit learning task.

The syntactic comprehension task we utilized is a direct replication of the task used by Misyak and Christiansen (2007) as the authors provided the materials used in the present study. Similarly, these authors collected their sentences from several previous authors, and combined them to create a measure of syntactic comprehension (Trueswell, Tanenhaus and Garnsey, 1994; Farmer, Christiansen and Monaghan, 2006; Wells et al., 2007).

The syntactic comprehension task presents participants with difficult to comprehend sentences, one-word at a time via the moving window paradigm (Just, Carpenter, and Woolley, 1982). Using this paradigm, the participant controls the movement of the window, thus revealing words at their discretion. Use of this paradigm in the context of the syntactic comprehension task, reveals the difficult to comprehend positions in a sentence, as participants are more likely to pause the movement of the window when they encounter processing difficulty. The precise nature of the comprehension difficulty is represented in this task by syntactic ambiguity. Syntactic ambiguity can represent the positions in a sentence that lead the reader to multiple interpretations, but the interpretation most intuitively derived is likely to be incorrect. The reasons for why ambiguity arises and how it is resolved is a highly complex issue, and beyond the scope of the present paper. It should suffice to indicate the sentences included in the syntactic comprehension task by Misyak and Christiansen (2007), and therefore the present study, are taken from several well-established paradigms of syntactic comprehension (Trueswell, Tanenhaus and Garnsey, 1994; Farmer, Christiansen and Monaghan, 2006; Wells et al., 2007). Examples of these sentences are presented in table 7.

Table 7. Syntactic Comprehension Task Examples

| |
|---|
| ***Animate/Inanimate Noun Clauses*** |
| Reduced: *The man/car towed by the garage was parked illegally.* |
| Unreduced: *The [man who]/[car that] was towed by the garage was parked illegally.* |
| ***Subject/Object Relative Clauses*** |
| Subject relative: *The player that noticed the coach threw the football across the field.* |
| Object relative: *The player that the coach noticed threw the football across the field.* |

It should be noted that as this task is an experimental measure of syntactic comprehension and thus has not been used normatively, it might not appear ideal for use in predicting clinically relevant individual differences in implicit learning. However, no clinical measure is available, to our knowledge, which focuses solely on specific linguistic constructs in adults at the specificity at which they do in children and adolescents. Thus, prior to obtaining a measure of individual difference on the syntactic comprehension task, the task was assessed to determine if the typical patterns observed in the literature were replicated in our study. That is, sentences were separately analyzed via the specific sentence type as animate forms were compared to inanimate forms across unreduced vs. reduced forms (2 X 2 factorial comparison). Similarly, performance on object-relative sentences was compared to subject-relative sentences. Sentences of different types were not compared (e.g., animate sentences were not compared to relative clauses).

Once the measure is obtained however, the learning rate from probability 0.33 (which, as previously indicated, appears to be our strongest measure of implicit learning) was compared against performance on the syntactic comprehension task, IQ-2 and WAIS-IV WMI, and then each factor was regressed on the other creating a correlation matrix.

### *6.1. Methods*

#### *6.1.1. Participants*

A subset of the participants recruited for the Serial Reaction Time Task study went on to complete the syntactic comprehension, WASI, and WAIS measures during the same testing session based upon whether or not English was their primary language. Participants that were not primary English speakers did not complete these measures as appropriate analysis of the results of these English-based tasks would be confounded by familiarity with the English language.

*6.1.2. Apparatus*

A Macintosh computer running OS X 10.6.8 (Snow Leopard) was used to present sentences during the syntactic comprehension task using the Linger software package. Participants were seated in front of the Surface®, Macintosh, or testing stimuli, with the center of the stimuli (or presentation device) placed at the participants midline within comfortable reaching distance (unless otherwise indicated by the standardized testing instrument instructions).

*6.1.3. Stimuli*

The specific sentences used include: 28 each of reduced and non-reduced sentences (56 in total) containing animate/inanimate noun constructions; 20 each of noun-like or verb-like homonyms (40 total) having typical or atypical resolutions serving as "filler trials"; and 40 each of subject or object relative clauses (80 total) (see Table 4 for example sentences from each sentence type).

*6.1.4. Procedure*

Participants began this portion of the experiment by completing the Vocabulary, Matrix Reasoning, and Block Design subtests of the Wechsler Abbreviated Scale of Intelligence (Wechsler, 1999; ~40 minutes), followed by the Digit Span and Arithmetic subtests from the Wechsler Adult Intelligence Scale, Fourth Edition (WAIS-IV, 2003; ~10 minutes). Participants then completed the syntactic judgment task (~25 minutes).

For the syntactic comprehension task, participants were presented one sentence at a time, and one word at a time, on a PC computer screen. As the task utilized the moving window paradigm (Just, Carpenter, and Woolley, 1982), participants proceeded to read the sentence one word at a time, moving themselves forward with the space bar, until the sentence was completed. As soon as the sentence was complete, the participant was asked a comprehension question requiring a yes or no response before proceeding to the next sentence. The same procedure was followed through the remaining 80 sentences.  Of these 80 sentences, eight were sentences with an animate noun as the subject of the sentences whereas another eight included an inanimate noun as the subject.  For each of these types, four were reduced and four were unreduced (allowing for a 2 X 2 comparison for these components) (please see table 4 for more details).  A further 20 of the sentences were subject/object relative sentences, with subject relative sentences representing sentence constructions in which the resolution of ambiguity of the sentence rests upon the subject of that sentence while the resolution of ambiguity rested upon the object of the sentence.  The remaining 54 sentence stimuli were considered the control, or “filler,” sentences.  For all the sentences, two versions were available for each, and random randomly assigned for each participant.  For example of each of the four animate/reduced sentences a participant will read, there are two options for each such that these sentences are randomly drawn from a total pool of eight.

To calculate reading times for each sentence, participant’s reading times of the sentences after the syntactic ambiguity was introduced into the sentence, were summed, dropping the first few words from the sentence prior to the ambiguity, from the analysis. This was done in order to focus analysis on the processing and comprehension from the introduction of the ambiguity onward. For example for the sentence "the player that noticed the coach threw the football across the field", the ambiguity is represented by the word "noticed" and therefore "the player that" was dropped from analysis, with "noticed the coach threw the football across the field" remaining for analysis.

*6.1.5. Clinical Instruments*

*6.1.5.1. Wechsler Intelligence Scales.*

The intellectual capabilities of young adults in our study, as represented by IQ, were measured with the two subtest form of full-scale IQ on the Wechsler Abbreviated Scale of Intelligence (WASI, IQ-2, Wechsler, 1999). Additionally, participants' working memory ability was measured using the Digit Span and Arithmetic subtests of the Working Memory Index within the Wechsler Adult Intelligence Scale-Fourth Edition (WAIS-IV, WMI; Wechsler, 2008). Stimuli and procedures used for the WASI and WAIS-IV adhered to standardized procedures (e.g., pen and paper recording, and required stimulus books).

*6.2. Statistical Modeling Approach*

A similar mixed-effects model building approach as that used for the serial reaction time task was utilized for the syntactic comprehension data to ensure they followed the general patterns represented in the literature. Further details about the modeling approach can thus be found in that section. However, it should be indicated that the base model from which the random effects were compared was defined as, "*Sentence Reading Speed* ~ 1 + (1|*participant*) + (1|*trial*)". It should additionally be noted that we did not compare sentences across types, meaning we did not compare performance on Animate Reduced trials to Subject-Relative clauses. Rather we broke the data into subsets and analyzed each sentence type (animate vs. inanimates together, subject vs. object relatives together) separately for all analyses. The single exception to this, is when data were cleaned, as the median sentence reading speed was based upon all sentences for the participant.

Upon finding a significant difference between sentence forms (animate/inanimate, reduced/unreduced) and clause types (subject/object relative), the individual difference components were extracted based upon performance of the easier model to the theoretically more difficult sentence (e.g., unreduced animate vs. reduced animate; object-relative vs. subject-relative) with performance on the more difficult model subtracted from the easy value, thereby giving a negative number under conditions in which the ambiguity resolution caused significant difficulty for the participant. These extracted values were then compared to the extracted value from the serial reaction time task ($P(0.33)$), as well as to working memory and IQ-2 using a simple linear regression without covariates, thus creating a correlation matrix.

*6.2.1. Random effects compared for models of syntactic comprehension performance.*

The random effects of *trial* and *participant* reflect similar random effects components as those compared in the serial reaction time task, allowing for variability associated with participant, and trial reflecting potential fatigue effects. Additionally, *base reading speed* was tested as a random effect and computed based upon performance of the control and filler sentences. *Sentence number* was included as it indicated the specific version of the randomly assigned sentence that the participant read. *Word count* was included as representing the number of words within a sentence.

*6.2.1. Fixed effects included in the model of syntactic comprehension performance.*

*Sentence type* was included in the model comparing Animate and Inanimate sentences, with inanimate sentences serving as the base contrast. The animate and inanimate sentences were further compared across the unreduced or reduced forms, labeled *Forms*, with unreduced form serving as the base level contrast. *Relative clause* represented either object or subject relative sentences with object-relatives serving as the base contrast.

*6.3. Results*

*6.3.1. Data cleaning*

Prior to data analysis, the median and standard deviation of these sums was then computed and values 2.5 standard deviations above the median were capped at 2.5

SD above the median for each participant. The data cleaning procedure used was thus similar to that used for the serial reaction time data for the pilot and primary experiments. This resulted in few trials requiring capping (Mean = 1.85% (1.22%), Range = [0%, 5%]). Working memory data was not available for two participants and so these participants were excluded from comparisons between working memory and other variables. Full-scale IQ-2 based upon the vocabulary and matrix reasoning subtests of the WASI was available for all of the participants whom performed the syntactic comprehension task.

*6.3.2. Modeling of Random Effects*

The results indicate that the specific sentence significantly improved model fit beyond the base model (BASE vs. M1; $\chi^2$ (1) = 104.2, $p < 0.001$) indicating that certain sentences were read faster than others. Inclusion of word count further improved model fit beyond the inclusion of sentence number (M1 vs. M2; $\chi^2$ (1) = 16. 652, $p < 0.001$), indicating that the number of words in the sentence also influenced reading speed. Further, including the individual's baseline reading speed in the model further improved model fit (M2 vs M3, $\chi^2$ (2) = 16.652, $p < 0.001$) indicating that during the control and filler sentences, some participants were faster than others.

Results were similar for the relative clauses in that sentence number and word count each improved model fit (in order: BASE vs M1, $\chi^2$ (1) = 104.2, $p < 0.001$; M1 vs. M2, $\chi^2$ (1) = 104.2, $p < 0.001$). However, baseline reading speed during the control sentences didn't influence average reading speed for the relative clauses (M2 vs M3, $\chi^2$ (1) = 0, p = 1.00).

BASE: *reading speed* ~ 1 + (1|*participant*) + (1|*trial*)

M1: BASE + (1|*sentence number*)

M2: M1 + (1|*word count*)

M3: M2 + (1|*baseline reading speed*)

*6.3.3. Modeling of Fixed Effects*

The results indicate that reading time of inanimate sentences did not differ significantly from animate forms compared to the base model (M3 vs. M4; $\chi^2$ (2) = 0.480, p = 0.787). Similarly, collapsing across the animate/inanimate distinction, reading times did also not differ between unreduced and reduced forms (M3 vs. M5; $\chi^2$ (1) = 2.098, p = 0.148). Modeling an interaction between animate/inanimate and reduced/unreduced forms also did not improve model fit beyond the base model (M3 vs M6; $\chi^2$ (6) = 2.572, p = 0.860). Upon visual inspection (see right panel of figure 7) the data appear to trend toward animate forms being the most difficult and showing very little difference between reduced and unreduced forms, while inanimate forms trend toward being read more quickly than animate sentences. The reduced form of the inanimate sentences appears to be read more slowly than the unreduced form. Nevertheless this trend does not approach significance. Thus for the animate and inanimate sentence distinction the model best fitting the data was revealed to be the final random effects model (M3).

M4: M3 + *Sentence Type* + (1|*participant*:*Sentence Type*)

M5: M3 + *Form* + (1|*participant*:*Form*).

M6: M4 + M5 + (*Sentence Type* * *Form*) + (1|*participant*:[Sentence Type * Form])

In comparing relative clauses, including the relative aspect of the sentences significantly improved model fit beyond the base model (M3 vs. M7; $\chi^2$ (2) = 16.659, p < 0.001).

**Table 8. Parameter Estimates of Logged Speed for the Fixed Effects Model of Object- vs. Subject-Relative Clauses**

| | Estimates | Std. Error | *t* | *p* |
|---|---|---|---|---|
| Intercept | 3.656 | 0.032 | 113.90 | 0.000 |
| Object vs. Subject Relative | 0.029 | 0.006 | 3.67 | 0.012 |

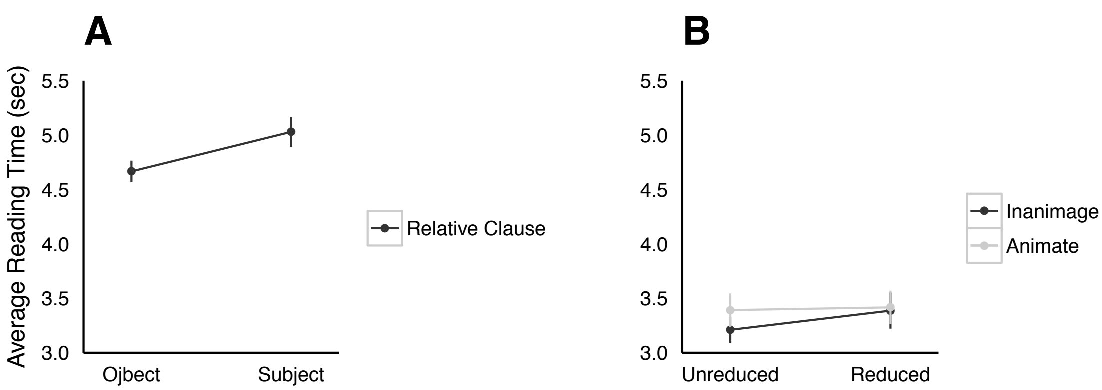


**Figure 7.** ***Results fom the syntactic Comprehension Task.*** Panel A indicates the results of the subject/object relative clauses, Results based upon these sentences indicate that subject-relative sentences were read significantly more slowly than object-relative sentences. Panel B represents the data from the sentences with animate/inanimate clauses and reduced or unreduced sentences. Results indicate no difference between animate or inanimate sentences, between reduced or unreduced sentences forms, or an interaction between the two.

The estimates, provided in table 8, indicate that subject relative sentences were read significantly slower than object relative sentences. Both the animate/inanimate data and the data from the relative clauses are shown in the left panel of figure 7.

M7: M3 + *Relative Clause* + (1|*participant*:*Relative Clause*)

The correlations indicate that performance on subject-object relative sentences were not correlated with any other performance measure utilized (WMI, IQ-2, or learning rate), indicating that these performance measures were not able to account for the differences observed between the disparate reading rates of subject- as compared to object-relative sentence forms. Similarly, learning rate as measured by the difference between block 1 and block 3 for probability 0.33 (our probability with the highest degree of learning) was not associated with WMI or IQ-2. Performance on the IQ-2 and WMI were unsurprisingly significantly correlated.

**Table 9. Correlation between Indices of Individual Differences in the Present Study**

| | Subj-Obj | WMI | IQ-2 | P(0.33):B1-B3 |
|---|---|---|---|---|
| Subj-Obj | . | . | . | . |
| WMI | 0.001 | . | . | . |
| IQ-2 | 0.000 | 0.697** | . | . |
| P(0.33):B1-B3 | -0.049 | 0.000 | 0.000 | . |

*Subj-Obj* indicates the difference score subtracting reading rate of object-relative from subject-relative sentences with higher values indicating slower reading speed of subject compared to object-relative sentences. *WMI* indicates the scaled score of the working memory index of the WAIS-IV. *IQ-2* represents the IQ score from the 2-subtest version of the WASI. *P(0.33):B1-B3* indicates our measure of implicit learning from the SRT task. Asterisks represent levels of significance for correlations, reported as * = $p$ < 0.05, * = p < 0.01, and *** as $p$ < 0.001.

*6.4. Discussion*

Overall, these data indicate that a reliable distinction between object and subject-relative sentences was found in our iteration of the syntactic comprehension task. We were thus able to utilize this as the component of individual differences, comparing performance on this task to performance on our implicit learning task as well as to working memory and full-scale IQ. The results suggest no relationship among any of the factors, except, unsurprisingly, WMI and full-scale IQ2. This is surprising as in previous literature, at minimum, Misyak and colleagues (2007, 2012) found an association between working memory and performance on subject/object relative sentences. Additionally they found a significant relationship between their implicit learning task and performance on the syntactic comprehension task above and beyond the influence of working memory. There are a number of interpretation one could take in explaining this lack of association, as our implicit learning task differed from theirs in a number of ways. Specifically, the implicit learning task they utilized in previous studies was a language based measure of implicit learning. Thus potentially providing for a heightened chance of observing a correlation between their tasks.

These data additionally indicate that our implicit learning task is not associated with working memory or full-scale IQ. This is unsurprising, and we were not particularly convinced that such an effect would be observed, as, in the introduction, we argued that implicit learning likely predicts other, but important, component processes not typically measured with cognitive batteries.

It is however surprising that we did not show a difference between the reduced and unreduced forms of animate and inanimate sentences as this is a fairly robust finding in the linguistics literature (cf. Misyak and Christiansen, 2012). However, fourth things should be considered regarding this interpretation. First, in the literature, the effects observed are typically, very very small, on the order of tens of milliseconds. It is notable, then, that even in the absence of significance, data visually tend to trend in the direction, commonly observed in the literature (i.e., reduced forms harder than unreduced, particularly for sentences containing an animate noun as the subject). Second, as this effect is robust but small, our very limited trials of 4 items per comparison per subject (i.e., recall the 2 X 2 factorial design) may have provided too few trials per subject upon which to compare performance. Third, in the research paradigms typically utilizing this task, they are not often contained in the same sentence sets as other ambiguous forms.

As these data reveal some consistency with the literature, one might then argue that we did not find any meaningful association between individual differences on our implicit learning task, and syntactic comprehension capability. We address this possibility in the general discussion, amidst other considerations concerning the shape and meaning of our data as a whole.

## Chapter 7: General Discussion

Throughout our study, we attempt to design and implement a modification to the well-known and widely used serial response time task with the goal of providing an easy to understand, implicitly motivating task that might be appropriately utilized as a clinical tool with a variety of clinical populations. Taken together, these results indicate that as in the pilot study, learning effects in our new SRT paradigm cannot be interpreted apart from consideration of movement related variables. Indeed, collapsing across all movement variables of the study doesn't show any significant increase or decrease in performance over time (as indicated by the parameter estimates of table 5, specifying the effect of block beyond the model including only random effects). However, when considering the task for what it truly is, that is, a task that is highly dependent upon the effects of movement, patterns of learning begin to emerge. As a significant boon to the argument that these movement factors must be considered, we found relatively consistent and significant effects for both the pilot and primary experiment, even though the screen size differences, and thus stimulus size differences, were wildly disparate. As indicated in the introduction, our secondary objective was to determine if learning on our implicit learning task was associated with performance on other higher-level cognitive tasks. Thus we require a measure of individual differences on our task to compare against these other tasks. However as previously stated, simply taking the difference between block 3 and block 4 from our experiment, as has been done in past studies (Brown et al., 2011) is inappropriate, as it completely neglects how learning occurs in our task. In order to specify our measure of individual differences we therefore carefully consider the specific patterns of performance for our task, and more specifically, the very influential patterns of movement. It is to this discussion we turn next, most carefully considering probability of movement, as it is here we see the greatest effect of learning.

Across both the pilot and primary experiments it appeared that the directional probability, specifying the probability of relative directions of the stimuli, significantly influenced performance, but only in specific circumstances. Indeed the "classic" v-shaped learning effect curve of the SRT task (i.e., speed increase across a predictable pattern, then a decrease in performance after the pattern is removed) was only observed for probability 0.33 for the participants of the primary experiment, while participants in the pilot experiment appeared to show the greatest effect for both 0.66, but not for 0.33. Nevertheless cutting across the different learning patterns for these lower probabilities, it was observed that the highest probability directional movement (1.00) was also associated with the slowest performance.

These data and patterns of significance can be well observed in figure 6. Trends in the data indicate that participants in the pilot study (left panel) as well as participants in the primary experiment (right panel) showed a clear indication of significantly lower speed when stimulus locations are in the direction of an absolute directional probability (1.00) then when the bubble to be popped has a lower directional probability (i.e., 0.33 and 0.66 having relatively equivocal performance at block 1). As stated when we introduced magnitude, each probability can only be fully compared at the lowest magnitude changes as, for example, the highest magnitudes necessitate an absolute probability of 1.00. Given this requirement, it is nevertheless important to consider the bottom panel of figure 5, breaking apart speed of performance by block, probability and magnitude across the trials. As previously stated, it is clear that participants are significantly faster in moving to higher magnitudes, and this is what is represented by the overall lower intercept at the highest magnitude. Nevertheless when considering movement probabilities that have equal magnitude (only possible from the "low" magnitudes), an absolute probability of 1.00 is associated with the slowest performance (as indicated by the significant patterns in the effects from the planned contrasts). This pattern holds at moderate magnitudes, in that the possible probabilities for comparison (0.66 and 1.00) reveal that again, probabilities of 1.00 are associated with significantly slower performance.

To further understand these data, the practice trials were also plotted (but never analyzed). By investigating the trend over practice trials, it can be seen that the greatest level of performance increase appears to occur during these initial trials in the practice block. It is during the practice block, and to some very small degree block 1, that participants may become quickly familiar with the strategies required to perform the task optimally. For example, it may be here that they come to first learn and utilize the broader rules governing the experiment. Specifically, prior to performance of the practice block, stimuli on our touch screen task could conceivably be presented at any location on the screen as is typical of touchscreen applications, games, etc. It is likely the case participants quickly come to understand that bubbles appear in one of four horizontal quadrants, even they are unsure of where the next trial would be. This attention allocation strategy is likely quickly utilized to better perform the experiment. Additionally, improvement in attentional allocation strategy would potentially come with increased awareness (either implicit or explicit) of directional probability. That is, once participants learn that from the furthest right location, they only have one direction to move, they are likely to begin utilizing this rule to help improve performance.

However, as revealed by the modeling of the data as well as through graphical depiction, it is clear that participants in fact are slower when there is an absolute probability of movement direction. Given our discussion to this point, these results are quite surprising. To use the whack-a-mole example, the assumption we have made is that from the bottom left most mole position in the whack-a-mole game, participants can be certain that they are going to be whacking the next mole in any position up or to the right of that position (assuming no repeats of mole position). They might thus prepare their arm and thus their mallet for a thrust in that direction. Given the results of our experiment however, what is actually suggested is that participants would be slower to whack the next mole if there is only a single direction they can move. Providing an alternative explanation requires for a rethinking of the situation the mole-whacker is exposed to.

To maintain the situation discussed thus far, our mole-whacker has just whacked the mole at the furthest bottom left position, leaving only mole holes up and to the right. Though from this position the person can be certain they will not be moving any further left, they are now presented with a problem: “which mole should be hit to the right?” Now that every stimulus position is up and to the right, they now actually have the largest number of choices for which they could prepare for, and thus the greatest degree of uncertainty in particular stimulus choice. To shift back to our task, even though from the furthest right position participants can be certain of the direction they will move left, from this furthest right positions there are now three different possible positions located to the left of which to choose. Thus absolute *directional probability* is associated with a far greater number of competitor objects to choose from, and thus specific item choice if far less certain. If in contrast, the participant is just to the right of center, they may only have a 33% chance of moving to the right, but if they were to move to the right they would have a 100% chance of accurately choosing the correct bubble should it appear on the right.

The competition amongst object choices having a negative impact on performance is a very well established effect in the experimental literature (c.f., Hommel, 2007). Specifically, it has been found on numerous occasions and among various paradigms, that the more possible choices an individual is presented with, the longer they will take to make their choice (Schiffer, 1998). Individual differences in this ability to choose quickly based upon previous instances of picking the same stimulus (as happened with our task) may continually reinforce the tendency to focus toward the 33% probability as the sequence might be easiest to understand if it is simply that sequence. Thus *choice probability* could play a significant role on multiple aspects of this task.

However, we are presently limited by a slightly higher signal-to-noise ratio in our detection of learning than we had originally anticipated. Specifically, we had originally opted to design a task that was as similar to the classic SRT as possible, even using a 12-item sequence that is one of the most widely utilized in the literature (Jimenez et al., 2005). Nevertheless, on the whole, for the primary experiment, we did not see a learning effect that was robust across all participants and situations (thus leading to an overall learning effect). Thus the meaningfulness of those individual differences may be so highly specific to a particular probability and to a particular sequence item, that the effects will not show transfer in the task.

In order to reduce the signal-to-noise ratio, we may be able to push all of the variables in our task toward learning. That is, now that we are aware of the noise that is present in our task, we may easily manipulate it through simple modifications. For example in order to reduce the effect of magnitude, it would likely be helpful to creating a “landing point” in the middle of a circular array of stimuli, that the participant must press between trials, prior to being shown the next trial. In this manner, magnitude could be careful controlled, as could probability (from a circular array, it would be simple to make items equally probable or push probability in the desired direction).

## Conclusion

Implicit learning is becoming recognized in several lines of literature as an ability to account for when predicting changes in performance, not readily attributable to explicit learning and memory processes. Significantly, in the very recent past, investigators have argued that implicit learning processes lie at the core of certain deficits such as autism spectrum disorders (Igetsi, 2013). In response to this pattern in the field, it is necessary that a normative tool be constructed that shows solid metrics of implicit learning, on that would be able to accurately measure individual differences. As we cannot conclude that our task hits the mark in this regard, given the consistency in our results across disparate paradigms, and the highly correlated pattern of individual differences in implicit learning across participants, we argue that we are well on our way, toward creating such a task.

**References**

Carroll JB. 1993. *Human Cognitive Abilities: A Survey of Factor Analytic Studies.* Cambridge, UK: Cambridge Univ. Press.

Carrow-Woolfolk, E. (1999). *Comprehensive Assessment of Spoken Language (CASL)*. Circle Pines, Minn, American Guidance Service.

Constantino, J. N. & Gruber, C. P. (2005). *Social Responsiveness Scale (SRS).* Western Psychological Services.

Courchesne, E., Carper, R., & Akshoomoff, N. (2003). Evidence of brain overgrowth in the first year of life in autism. *JAMA: the journal of the American Medical Association*, *290*(3), 337–344.

Conway, C.M. & Christiansen, M.H. (2006). Statistical learning within and between modalities: Pitting abstract against stimulus specific representations. *Psychological Science*, 17, 905-912.

Danner, D., Hagemann, D., Schankin, A., Hager, M., & Funke, J. (2011). Beyond IQ: A latent state-trait analysis of general intelligence, dynamic decision making, and implicit learning. *Intelligence*, *39*(5), 323–334. doi:10.1016/j.intell.2011.06.004

Destrebecqz, A., & Cleeremans, A. (2001). Can sequence learning be implicit? New evidence with the process dissociation procedure. *Psychonomic Bulletin & Review*, *8*(2), 343–350.

Eigsti, I., & Mayo, J. (2011). Implicit learning in ASD. In D. A. Fein (Ed.), *The neuropsychology of autism.* New York: Oxford Press.

Evans, J. L., Saffran, J. R., & Robe-Torres, K. (2009). Statistical learning in children with specific language impairment. Journal of Speech, *Language, and Hearing Research*, *52*(2), 321.

Folia, V., Uddén, J., Forkstam, C., Ingvar, M., Hagoort, P., & Petersson, K. M. (2008). Implicit learning and Dyslexia. *Annals of the New York Academy of Sciences*, *1145*(1), 132–150. doi:10.1196/annals.1416.012

Farmer, T. A., Christiansen, M. H., & Monaghan, P. (2006). Phonological typicality influences on-line sentence comprehension. *Proceedings of the National Academy of Sciences*, *103*, 12203-12208.

Funke, J. (1983). Einige Bemerkungen zu Problemen der Problemlöseforschung oder: Ist Testintelligenz doch ein Prädiktor? *Diagnostica, 29*, 283–302.

Gebauer, G. F., & Mackintosh, N. J. (2007). Psychometric intelligence dissociates implicit and explicit learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(1), 34–54. doi:10.1037/0278-7393.33.1.34

Geschwind, D. H. (2009). Advances in Autism. *Annual Review of Medicine*, *60*(1), 367–380. doi:10.1146/annurev.med.60.053107.121225

Geschwind, D. H., & Levitt, P. (2007). Autism spectrum disorders: developmental disconnection syndromes. *Current Opinion in Neurobiology*, *17*(1), 103–111. doi:10.1016/j.conb.2007.01.009

Gibson, J. J., & Gibson, E. J. (1955). Perceptual learning: Differentiation or enrichment? *Psychological Review*, *62*(1), 32.

Happé, F., Ronald, A., & Plomin, R. (2006). Time to give up on a single explanation for autism. *Nature Neuroscience*, *9*(10), 1218–1220. doi:10.1038/nn1770

Hyman, L.M. (1977). On the nature of linguistic stress. In L. M. Hyman (Ed.), *Studies in stress and accent.* Southern California Occasional Papers in Linguistics, No. 4. Los Angeles: University of Southern California.

Jäger, A. O., Süß, H. M., & Beauducel, A. (1997). *Berliner Intelligenzstruktur — Test. Form 4.* Göttingen: Hogrefe.

Jiménez, L., & Vázquez, G. A. (2005). Sequence learning under dual-task conditions: alternatives to a resource-based account. *Psychological Research Psychologische Forschung*, *69*(5-6), 352–368. doi:10.1007/s00426-004-0210-9

Just, M. A., Carpenter, P. A., & Woolley, J. D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General*, *111*(2), 228.

Kaufman, S. B., DeYoung, C. G., Gray, J. R., Jiménez, L., Brown, J., & Mackintosh, N. (2010). Implicit learning as an ability. *Cognition*, *116*(3), 321–340. doi:10.1016/j.cognition.2010.05.011

Kessels, R. P. C., Remmerswaal, M., & Wilson, B. A. (2011). assessment of nondeclarative learning in severe Alzheimer dementia: the Implicit Memory Test (IMT). *Alzheimer Disease & Associated Disorders*, *25*(2), 179.

Kidd, E. (2012). Implicit statistical learning is directly associated with the acquisition of syntax. *Developmental Psychology*, *48*(1), 171–184. doi:10.1037/a0025405

Klinger, L. G., & Dawson, G. (2001). Prototype formation in autism. *Development and Psychopathology*, *13*(01), 111–124.

Klinger, L. G., Klinger, M. R., & Pohlig, R. A. (2007). Implicit learning impairments in autism spectrum disorders: Implications for treatment. In J. M. Perez, P. M. Gonzalez, M. L. Comi, & C. Nieto (eds.), *New developments in autism*. London: Jessica Kinglsey.

Knowlton, B. J., Mangels, J. A., & Squire, L. R. (1996). A neostriatal habit learning system in humans. *Science*, *273*(5280), 1399–1402.

Lieberman, M. D. (2000). Intuition: A social cognitive neuroscience approach. *Psychological Bulletin*, *126*(1), 109–137. doi:10.1037//0033-2909.126.1.109

Lord, C, Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B. L., DiLavore, P. C., . . .Rutter, M. (2000). The Autism Diagnostic Observa- tion Schedule-Generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal ofAutism and Developmental Disorders*, *30*, 205-223.

Lyons, J. (1968). *Introduction to Theoretical Linguistics*. Cambridge: Cambridge University Press.

Mackintosh, N. J. (1998). *IQ and human intelligence.* Oxford: Oxford University Press.

Misyak, J. B., & Christiansen, M. H. (2007). Extending statistical learning farther and further: Long-distance dependencies, and individual differences in statistical learning and language. *Proceedings of the 29th Annual Cognitive Science Society*, 1307–1312.

Misyak, J. (2010). On-line individual differences in statistical learning predict language processing. *Frontiers in Psychology*, *1*. doi:10.3389/fpsyg.2010.00031

Misyak, J. B., Christiansen, M. H., & Bruce Tomblin, J. (2010). Sequential Expectations: The Role of Prediction-Based Learning in Language. *Topics in Cognitive Science*, *2*(1), 138–153. doi:10.1111/j.1756-8765.2009.01072.x

Misyak, J.B., Goldstein, M.H. & Christiansen, M.H. (in press). Statistical-sequential learning in development. In P. Rebuschat & J. Williams (Eds.), *Statistical learning and language acquisition.* Berlin: Mouton de Gruyter

Mitchell, C. J., De Houwer, J., & Lovibond, P. F. (2009). The propositional nature of human associative learning. *Behavioral and Brain Sciences*, *32*(02), 183. doi:10.1017/S0140525X09000855

Mostofsky, S. H., Goldberg, M. C., Landa, R. J., & Denckla, M. B. (2000). Evidence for a deficit in procedural learning in children and adolescents with autism:

implications for cerebellar contribution. *Journal of the International Neuropsychological Society*, *6*(7), 752–759.

Müller, R.-A., Cauich, C., Rubio, M. A., Mizuno, A., & Courchesne, E. (2004). Abnormal activity patterns in premotor cortex during sequence learning in autistic patients. *Biological Psychiatry*, *56*(5), 323–332. doi:10.1016/j.biopsych.2004.06.007

Newport, E. L. (2010). Plus or minus 30 years in the Language Sciences. *Topics in Cognitive Science*, *2*(3), 367–373. doi:10.1111/j.1756-8765.2010.01102.x

Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology*, *19*(1), 1–32.

Ozonoff, S., Goodlin-Jones, B. L., & Solomon, M. (2005). Evidence-based assessment of autism spectrum disorders in children and adolescents. *Journal of Clinical Child & Adolescent Psychology*, *34*(3), 523–540.

Pelucchi, B., Hay, J. F., & Saffran, J. R. (2009). Statistical Learning in a Natural Language by 8-Month-Old Infants. *Child development*, *80*(3), 674–685.

Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: one phenomenon, two approaches. *Trends in Cognitive Sciences*, *10*(5), 233–238. doi:10.1016/j.tics.2006.03.006

Pothos, E. M. (2007). Theories of artificial grammar learning. *Psychological Bulletin*, *133*(2), 227–244. doi:10.1037/0033-2909.133.2.227

Ratcliff, R., & Murdock, B. B., Jr. (1976). Retrieval processes in recognition memory. *Psychological Review*, 83, 190-214.

Raven, J. C., Court, J. H., & Raven, J. (1994). *Manual for Raven's progressive matrices and mill hill vocabulary scales. Advanced progressive matrices.* Oxford: Oxford Psychologists Press.

Reber, A. S. (1967). Implicit learning of Artificial Grammars, *Journal of Verbal Learning and Verbal Behavior*, *6*(6) p. 855 - 863.

Reber, A. S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General*, *118*(3), 219.

Reber, A. S., Walkenfeld, F. F., & Hernstadt, R. (1991). Implicit and explicit learning: Individual differences and IQ. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*(5), 888.

Saffran, J. (2002). Constraints on Statistical Language Learning. *Journal of Memory and Language*, *47*(1), 172–196. doi:10.1006/jmla.2001.2839

Shanks, D. R. 2005. Implicit learning. In K. Lamberts, R. Goldstone (Eds.), *Handbook of Cognition.* Sage, London, 202–220.

Sopena S, Dewar BK, Greenfield E,& Wilson, B.A. (2005). Implicit memory assessment and its implications for rehabilitation. *Brain Impair*ment, *6*(2), 143–144.

Travers, B. G., Klinger, M. R., Mussey, J. L., & Klinger, L. G. (2010). Motor-linked implicit learning in persons with autism spectrum disorders. *Autism Research*, *3*(2), 68–77. doi:10.1002/aur.123

Trueswell, J., & Tanenhaus, M. (1994). Toward a lexical framework of constraint-based syntactic ambiguity resolution. *Perspectives on sentence processing*, 155–179.

Tsatsanis, K. D., Dartnall, N., Cicchetti, D., Sparrow, S. S., Klin, A., & Volkmar, F. R. (2003). Concurrent validity and classification accuracy of the Leiter and Leiter-R in low-functioning children with autism. *Journal of Autism and Developmental Disorders*, *33*(1), 23–30.

Walenski, M., Mostofsky, S. H., Gidley-Larson, J. C., & Ullman, M. T. (2008). Brief report: enhanced picture naming in autism. *Journal of Autism and Developmental Disorders*, *38*(7), 1395–1399. doi:10.1007/s10803-007-0513-y

Walenski, M., Tager-Flusberg, H., & Ullman, M. T. (2006). Language in Autism. In Moldin, S. O., & Rubenstein, J. L. R. (Eds.), *Understanding Autism: From Basic Neuroscience to Treatment*. CRC Press.

Warrington, E. K., & Weiskrantz, L. (1974). The effect of prior learning on subsequent retention in amnesic patients. *Neuropsychologia*, *12*(4), 419–428.

Wechsler D. 1999. *Wechsler Abbreviated Scales of Intelligence (WASI).* San Antonio, TX: Psychological Corporation.

Wechsler, D. (2008). *Wechsler Adult Intelligence Scale—Fourth Edition (WAIS-IV)*. San Antonio, TX: Pearson Assessment.

Wells, J. B., Christiansen, M. H., Race, D. S., Acheson, D. J., & MacDonald, M. C. (2009). Experience and sentence processing: Statistical learning and relative clause comprehension. *Cognitive Psychology*, *58*(2), 250–271. doi:10.1016/j.cogpsych.2008.08.002

Wirth, J., & Klieme, E. (2003). Computer-based assessment of problem solving competence. *Assessment in Education: Principles, Policy & Practice*, *10,* 329–345. doi:10.1080/0969594032000148172.

Woodward, J.Z. & Aslin, R.N. (1990). Segmentation cues in maternal speech to infants. Poster presented at the International Conference on Infancy Studies, Montreal, Canada.